



PS Claim 2; Page; 70pp; English.  
XX  
CC AA07180-83 represent N-terminally truncated human lactoferrin proteins.  
CC The lactoferrin variants bind heparin with lower affinity than natural  
CC lactoferrin. The lactoferrin variants exhibit biological activities such  
CC as anti-inflammatory, anti-viral and anti-infective activities as well as  
CC a pro- and anti-coagulant effects, modulation of complement activation,  
CC inhibition of LPS mediated activation of neutrophils, inhibition of  
CC myeloperoxidase, regulation of transcription, growth promotion of intestinal  
CC epithelial cells, inhibition of hydroxyl-radical formation, and a role in  
CC intestinal iron uptake and excretion. note: this sequence does not appear  
CC in the specification; it was created using information provided  
XX  
XX Sequence 690 AA:  
QY 1 ARARVVMAAVG 12  
DB 339 ARARVVMAVAG 350  
Query Match 93.3%; Score 56; DB 2; Length 690;  
Best Local Similarity 91.7%; Pred. No. 0.24;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
RESULT 11  
AA077908  
ID AAG77908 standard; protein; 690 AA.  
XX  
AC AAG77908;  
XX  
DT 22-JAN-2002 (first entry)  
XX  
DE Human lactoferrin variant hLF-2N.  
XX  
XX Human; lactoferrin; hLF; infectious disease; inflammatory disease;  
XX excess of heparin; gastroenteritis; inflammatory bowel disease; sepsis;  
XX anaemia; myelopoiesis; reperfusion injury; cytokine release;  
XX proteoglycan; hLF-2N; mutant; mutain.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT Domain 26..29  
FT /label= Cationic\_domain  
XX  
PN WO200172322-A2.  
XX  
XX 04-OCT-2001.  
XX  
PD 27-MAR-2001; 2001WO-NL000253.  
XX  
PF 27-MAR-2001; 2000EP-00201110.  
XX  
PR 27-MAR-2000; 2000US-0193352P.  
XX  
XX (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.  
XX  
XX Van Bree JBMW, Nuijens JH;  
XX  
XX WPI; 2001-648424/74.  
XX  
DR WPI; 2001-648424/74.  
XX  
XX Use of lactoferrin for treatment of infectious diseases, inflammatory  
XX diseases and excess of heparin.  
XX  
XX Claim 10; Page; 49pp; English.  
XX  
CC The sequence represents the human lactoferrin (hLF) protein variant hLF-  
CC 2N. The invention relates to novel methods of treatment using high doses  
CC of lactoferrin. The methods of the invention are useful for the treatment  
CC or prophylaxis of infectious diseases, inflammatory diseases and excess  
CC of heparin e.g. gastroenteritis, inflammatory bowel disease, sepsis,  
CC anaemia, myelopoiesis, reducing reperfusion injury, cytokine release and  
CC proteoglycan-mediated entry of virus into cells. The advantage of the  
CC method is that the patient is substantially free of side effect responses

CC to administration of lactoferrin. Therefore large doses of lactoferrin  
CC can be administered. Note: The present sequence is not shown in the  
CC specification but is derived from human wild-type lactoferrin sequence  
CC given in the sequence listing of the specification  
XX  
XX Sequence 690 AA:  
QY 1 ARARVVMAAVG 12  
DB 339 ARARVVMAVAG 350  
Query Match 93.3%; Score 56; DB 4; Length 690;  
Best Local Similarity 91.7%; Pred. No. 0.24;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
RESULT 12  
AB080724  
ID AB080724 standard; protein; 690 AA.  
XX  
AC AB080724;  
XX  
DT 29-NOV-2002 (first entry)  
XX  
DE Codon optimised lactoferrin protein.  
XX  
XX Transformed plant; heterologous transcription factor; transgenic plant;  
XX seed protein; protein expression.  
XX  
XX Homo sapiens.  
XX  
OS WO200264750-A2.  
XX  
PN WO200264750-A2.  
XX  
PD 22-AUG-2002.  
XX  
XX 14-FEB-2002; 2002WO-US004909.  
XX  
PF 14-FEB-2001; 2001US-0269188P.  
XX  
PR 14-FEB-2001; 2001US-0269199P.  
XX  
PR 02-MAY-2001; 2001US-00847232.  
XX  
XX (VENT-) VENTRIA BIOSCIENCE.  
XX  
XX Huang N, Yang D;  
XX  
PI WPI; 2002-657592/70.  
XX  
DR N-PSDB; ABS66515.  
XX  
XX Producing heterologous polypeptide in plant grain, by culturing  
XX transformed plant to form a grain-producing transforming plant, and  
XX recovering transgenic grains containing polypeptide from transformed  
XX plant.  
XX  
XX Exemple 15; Page 117; 230pp; English.  
XX  
XX The invention describes a method of producing a heterologous polypeptide  
XX (I) in a grain of a plant, comprising culturing a transformed plant (P1)  
XX comprising a first chimeric gene, and optionally, at least one  
XX heterologous transcription factor that is capable of enhancing the  
XX expression of the chimeric gene, to form a grain producing transforming  
XX plant (P2), and recovering transgenic grains containing (I) from P2. The  
XX method is useful for producing heterologous polypeptide in a grain of a  
XX plant. (I) is a non-plant storage, human or non-human animal, milk or  
XX other than a milk polypeptide, antibodies, cytokines, lymphokines,  
XX chemokines, hormones, growth factors, coagulation factors, anti-  
XX infectives, or cytotoxins, or anti-inflammatory molecule or intestinal  
XX trefoil factor (ITF) or its active fragment. Preferably, (I) is  
XX lactoferrin, lysozyme, lactoferricin, ITF, epidermal growth factor (EGF),  
XX keratinocyte growth factor (KGF), insulin-like growth factor I (IGF-I),  
XX lactonectin, kappa-casein, heparin, lactoperoxidase, alpha-1-  
XX antitrypsin, immunoglobulins, alpha-lactalbumin, beta-lactoglobulin,  
XX alpha-casein, beta-casein, albumin, fibrinogen or protease inhibitor.  
XX This is the amino acid sequence of a protein associated with method of

QY 1 ARRARVWMAVG 12

neutrophil degranulation; myelopoiesis inhibition

```

OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..27
FT /note= "Fragments of the N-terminal are specifically
FT referred to in the claims"
FT 2..5
FT Domain /label= Cationic_domain
FT 28..31
FT Domain /label= Cationic_domain
XX
XX WO200134641-A2.
XX
XX 17-MAY-2001.
XX
XX 10-NOV-2000; 2000WO-NL000821.
XX
XX 11-NOV-1999; 99EP-00203775.
XX PR 11-NOV-1999; 99US-0164975P.
XX
XX (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.
XX
XX Van Berkel PHC, Nibbering PH, Nuijens JH;
XX
XX WFI; 2001-335909/35.
XX
XX New polypeptides comprising the N-terminal region of human lactoferrin
XX protein exhibit higher antimicrobial activity than the full length
XX protein and are useful to treat bacterial infections.
XX
XX Claim 1; Page 55-57; 59pp; English.
XX
XX This invention relates to fragments (between 6 and 26 amino acids) of the
XX human lactoferrin hLF protein (represented by the present sequence). N-
XX terminal hLF peptides have antimicrobial activity. The peptides of the
XX invention are used to treat microbial infections, especially infections
XX by gram positive or negative bacteria, particularly Listeria,
XX Staphylococcus, Klebsiella or Escherichia species, especially L.
XX monocytogenes, S. aureus and E. coli. Other uses include reducing
XX inflammatory response by neutralising heparin or lipopolysaccharide or by
XX reducing cytokine production and neutrophil degranulation, inhibiting
XX entry of viruses such as cytomegalovirus, herpes simplex virus 1 or HIV
XX into cells, inhibiting myelopoiesis and reducing production of GM-CSF
XX (granulocyte/macrophage colony stimulating factor)
XX
XX Sequence 692 AA.
XX
XX
XX Query Match 93.3%; Score 56; DB 4; Length 692;
XX Best Local Similarity 91.7%; Pred. No. 0.24;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 ARRAVWMAVG 12
XX |||||
XX 341 ARRAVWCAVG 352

```

Search completed: September 1, 2004, 00:09:50  
 Job time : 29.88 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2004, 00:04:10 ; Search time 7.76 Seconds

(without alignments)  
79,834 Million cell updates/sec

Title: US-09-508-095-22

Sequence: 1 ARRRARYMAVG 12

Scoring table: ELOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfillseq1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	93.3	694	3	US-08-724-586-2
2	56	93.3	694	3	US-09-421-632-2
3	56	93.3	694	4	US-09-932-190-2
4	56	93.3	705	2	US-08-655-640-2
5	56	93.3	708	1	US-08-655-640-4
6	56	93.3	709	2	US-08-154-019-2
7	56	93.3	709	1	US-08-461-333-2
8	56	93.3	709	3	US-08-464-167-2
9	56	93.3	709	3	US-09-158-313-2
10	56	93.3	709	3	US-08-476-798-2
11	56	93.3	711	1	US-08-145-681-2
12	56	93.3	711	1	US-08-250-308-2
13	56	93.3	711	1	US-08-154-019-4
14	56	93.3	711	1	US-08-461-333-4
15	56	93.3	711	1	US-08-453-703-2
16	56	93.3	711	2	US-08-456-106-2
17	56	93.3	711	3	US-08-464-167-4
18	56	93.3	711	3	US-09-158-313-4
19	56	93.3	711	3	US-08-456-108-2
20	56	93.3	711	3	US-08-476-798-4
21	56	93.3	711	3	US-09-265-577-2
22	56	93.3	711	4	US-09-633-739-2
23	56	93.3	711	5	PCT-US93-03614-2
24	49	81.7	703	1	US-08-145-681-6
25	49	81.7	703	1	US-08-453-703-6
26	49	81.7	703	2	US-08-456-106-6
27	49	81.7	703	3	US-08-456-108-6

28	49	81.7	703	3	US-09-265-577-6	Sequence 6, Appli
29	49	81.7	703	4	US-09-633-739-6	Sequence 6, Appli
30	45	75.0	708	1	US-08-145-681-4	Sequence 4, Appli
31	45	75.0	708	1	US-08-453-703-4	Sequence 4, Appli
32	45	75.0	708	2	US-08-456-106-4	Sequence 4, Appli
33	45	75.0	708	3	US-08-456-108-4	Sequence 4, Appli
34	45	75.0	708	3	US-09-265-577-4	Sequence 4, Appli
35	45	75.0	708	3	US-09-633-739-4	Sequence 4, Appli
36	43	71.7	884	4	US-09-252-991A-26707	Sequence 2, Appli
37	40	66.7	135	4	US-09-252-991A-23456	Sequence 2, Appli
38	40	66.7	328	2	US-09-977-767-1	Sequence 1, Appli
39	39	65.0	290	4	US-09-252-991A-31211	Sequence 1, Appli
40	38	63.3	284	1	US-08-061-889-2	Sequence 2, Appli
41	38	63.3	284	1	US-08-462-611-2	Sequence 2, Appli
42	38	63.3	284	4	US-08-623-428D-2	Sequence 2, Appli
43	38	63.3	284	5	PCT-US94-05378-2	Sequence 2, Appli
44	38	63.3	379	4	US-09-252-991A-26689	Sequence 2, Appli
45	38	63.3	449	4	US-09-134-000C-4588	Sequence 4588, Ap

#### ALIGNMENTS

RESULT 1  
US-08-724-586-2  
Sequence 2, Application US/08724586  
Patent No. 6066469  
GENERAL INFORMATION:  
APPLICANT: Krugel, Marian L.  
APPLICANT: Kurecki, Tomasz  
APPLICANT: Gollnick, Paul D.  
APPLICANT: Doyle, Darrell U.  
TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
TITLE OF INVENTION: Lactoferrin  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jacobson, Price, Holman & Stern  
STREET: 400 Seventh St. N.W.  
CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,586  
FILING DATE: 30-SEPT-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/238,445  
FILING DATE: 05-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: 10505/P58185C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-724-586-2  
Query Match 93.3%; Score 56; DB 3; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.1;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 343 ARARVWCAVG 354

## RESULT 2

US-09-421-632-2  
Sequence 2, Application US/09421632

Patent No. 6277817

## GENERAL INFORMATION:

APPLICANT: Kruzel, Marian L.

APPLICANT: Kruzel, Marian L.

APPLICANT: Golnick, Paul D.

APPLICANT: Doyle, Darrell J.

TITLE OF INVENTION: Cloning, Expression, and Uses of Human

TITLE OF INVENTION: Lactoferrin

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Jacobson, Price, Holman & Stern

STREET: 400 Seventh St. N.W.

CITY: Washington D.C.

COUNTRY: U.S.A.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/421,632

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/724,586

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Player, William E.

REGISTRATION NUMBER: 31,409

REFERENCE/DOCKET NUMBER: 10505/P58185C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-6666

TELEFAX: (202) 393-5350

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 694 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-421-632-2

Query Match 93.3%; Score 56; DB 3; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.1;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ARARVWCAVG 12  
Db 343 ARARVWCAVG 354

## RESULT 3

US-09-932-190-2

Sequence 2, Application US/09932190

Patent No. 6455687

## GENERAL INFORMATION:

APPLICANT: Kruzel, Marian L.

APPLICANT: Kruzel, Marian L.

APPLICANT: Golnick, Paul D.

APPLICANT: Doyle, Darrell J.

TITLE OF INVENTION: Cloning, Expression, and Uses of Human

TITLE OF INVENTION: Lactoferrin

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Jacobson, Price, Holman & Stern

STREET: 400 Seventh St. N.W.

CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/932,190

FILING DATE: 17-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/724,586

FILING DATE: 30-SEPT-1996

APPLICATION NUMBER: US 08/238,445

FILING DATE: 05-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: Player, William E.

REGISTRATION NUMBER: 31,409

REFERENCE/DOCKET NUMBER: 10505/P58185C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-6666

TELEFAX: (202) 393-5350

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 694 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-932-190-2

Query Match 93.3%; Score 56; DB 4; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.1;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ARARVWCAVG 12  
Db 343 ARARVWCAVG 354

## RESULT 4

US-08-655-640-2

Sequence 2, Application US/08655640

Patent No. 5948613

## GENERAL INFORMATION:

APPLICANT: Teng, Christina

APPLICANT: Panella, Timothy J.

TITLE OF INVENTION: HUMAN LACTOFERRIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CUSHMAN, DARBY & CUSHMAN

STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/655,640

FILING DATE: 30-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/992,538

FILING DATE: December 17, 1992

ATTORNEY/AGENT INFORMATION:

NAME: SCOTT, WATSON T.

REGISTRATION NUMBER: 26,581

REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 705 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-655-640-2

Query Match 93.3%; Score 56; DB 2; Length 705;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
|||||  
Db 358 ARRAVWMAVG 369

RESULT 5  
US-08-655-640-4  
Sequence 4, Application US/08655640  
Patent No. 5948613  
GENERAL INFORMATION:  
APPLICANT: Teng, Christina  
APPLICANT: Patella, Timothy J.  
TITLE OF INVENTION: HUMAN LACTOFERRIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/655,640  
FILING DATE: 30-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/992,538  
FILING DATE: December 17, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: SCOTT, WATSON T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 708 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-655-640-4

Query Match 93.3%; Score 56; DB 2; Length 708;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
|||||  
Db 360 ARRAVWMAVG 371

RESULT 6  
US-08-154-019-2  
Sequence 2, Application US/08154019  
Patent No. 5633076  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heymaker, Herbert L.  
APPLICANT: Plateburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krumpfenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/154,019  
FILING DATE: 16-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/461,333  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebescheut, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-545-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-154-019-2

Query Match 93.3%; Score 56; DB 1; Length 709;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
|||||  
Db 358 ARRAVWMAVG 369

RESULT 7  
US-08-461-333-2

Sequence 2, Application US/0846133  
Patent No. 5741957  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,333  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschuetz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-461-333-2

Query Match. 93.3%; Score 56; DB 1; Length 709;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVYMAVVG 12  
DB 358 ARARVYMAVVG 369

RESULT 8  
US-08-464-167-2  
Sequence 2, Application US/08461167  
Patent No. 6013857  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.

APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,167  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschuetz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003124  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-167-2

Query Match. 93.3%; Score 56; DB 3; Length 709;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVYMAVVG 12  
DB 358 ARARVYMAVVG 369

RESULT 9  
US-09-158-313-2  
Sequence 2, Application US/09158313  
Patent No. 6067725  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krumpfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods

```

NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/158,313
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,798
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/895,956
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/619,131
FILING DATE: 27-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 709 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-158-313-2

Query Match          93.3%; Score 56; DB 3; Length 709;
Best Local Similarity 91.7%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ARRARVYMAVG 12
Db      358 ARRARVYMAVG 369

RESULT 10
US-08-476-798-2
Sequence 2, Application US/08476798
Patent No. 6140552
GENERAL INFORMATION:
APPLICANT: Deboer, Herman A.
APPLICANT: Strijker, Rein
APPLICANT: Heyneker, Herbert L.
APPLICANT: Platenburg, Gerald
APPLICANT: Lee, Sang He
APPLICANT: Pieper, Frank
APPLICANT: Krimpenfort, Paul J.A.
TITLE OF INVENTION: Production of Recombinant Polypeptides
TITLE OF INVENTION: by Bovine Species and Transgenic Methods
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
```

```

COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,798
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,788
FILING DATE: 15-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/895,956
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/619,131
FILING DATE: 27-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 709 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-798-2

Query Match          93.3%; Score 56; DB 3; Length 709;
Best Local Similarity 91.7%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ARRARVYMAVG 12
Db      358 ARRARVYMAVG 369

RESULT 11
US-08-145-681-2
Sequence 2, Application US/08145681
Patent No. 5571691
GENERAL INFORMATION:
APPLICANT: Conneely, Orla M.
APPLICANT: Heaton, Denis R.
APPLICANT: O'Malley, Bert W.
APPLICANT: May, Gregory S.
TITLE OF INVENTION: Production of Recombinant Lactoferrin
TITLE OF INVENTION: and Lactoferrin Polypeptides Using cDNA Sequences in
TITLE OF INVENTION: Various Organisms
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Baker & Botts, L.L.P.
STREET: 910 Louisiana St
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/145,681  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McGregor, Martin L.  
REGISTRATION NUMBER: 29,329  
REFERENCE/DOCKET NUMBER: 19928-0125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713/229/1874  
TELEFAX: 713/229/1522  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: H. sapiens  
US-08-145-681-2

Query Match 93.3%; Score 56; DB 1; Length 711;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 ARARVYMAVG 12  
DB 360 ARARVYMAVG 371

RESULT 12  
US-08-250-308-2  
Sequence 2, Application US/08250308.  
Patent No. 5571896  
Patent No. 5571896 5571896  
GENERAL INFORMATION:  
APPLICANT: Conneely, Orla M.  
APPLICANT: Haddon, Denis R.  
APPLICANT: O'Malley, Bert W.  
TITLE OF INVENTION: Production of Recombinant Human  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski Patent Department  
STREET: 1301 McKinney St.  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77010-3095  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/250,308  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/873,304  
FILING DATE: 24-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler Ph.D., Benjamin A.  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5456  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713/6515587  
TELEFAX: 713/6515246  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-250-308-2

Query Match 93.3%; Score 56; DB 1; Length 711;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 ARARVYMAVG 12  
DB 360 ARARVYMAVG 371

RESULT 13  
US-08-154-019-4  
Sequence 4, Application US/08154019  
Patent No. 5633076  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/154,019  
FILING DATE: 16-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/461,333  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschultz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-154-019-4

Wed Sep 1 08:24:13 2004

us-09-508-095-22.aug31.ra1

Page 7

Query Match 93.3%; Score 56; DB 1; Length 711;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVWMAVG 12  
DB 360 ARARVWCAVG 371

RESULT 14  
US-08-461-333-4  
Sequence 4, Application US/08461333  
Patent No. 5741957  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,333  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebescheutz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-461-333-4

Query Match 93.3%; Score 56; DB 1; Length 711;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVWMAVG 12  
DB 360 ARARVWCAVG 371

DB 360 ARARVWCAVG 371

RESULT 15  
US-08-453-703-2  
Sequence 2, Application US/08453703  
Patent No. 576939  
GENERAL INFORMATION:  
APPLICANT: Conneely, Orla M.  
APPLICANT: Headon, Denis R.  
APPLICANT: O'Malley, Bert W.  
APPLICANT: May, Gregory S.  
TITLE OF INVENTION: Production of Recombinant Lactoferrin  
TITLE OF INVENTION: and Lactoferrin Polypeptides Using cDNA Sequences in  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,703  
FILING DATE: Concurrently herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/145,681  
FILING DATE: October 28, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Albert P. Halluin  
REGISTRATION NUMBER: 25,227  
REFERENCE/DOCKET NUMBER: 8206-024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-3660  
TELEFAX: 415-854-3694  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: H. sapiens  
US-08-453-703-2

Query Match 93.3%; Score 56; DB 1; Length 711;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVWMAVG 12  
DB 360 ARARVWCAVG 371

Search completed: September 1, 2004, 00:18:41  
Job time: 8.76 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2004, 00:15:50 ; Search time 27.36 seconds  
(without alignments)

137.988 Million cell updates/sec

Title: US-09-508-095-22

Perfect score: 60

Sequence: 1 ARRARYVMAVG 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US06\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	93.3	359	US-10-169-297-49	Sequence 49, Appl
2	56	93.3	690	US-10-076-816-4	Sequence 4, Appl
3	56	93.3	690	US-10-077-81-4	Sequence 4, Appl
4	56	93.3	690	US-10-639-835-4	Sequence 4, Appl
5	56	93.3	694	US-10-023-096-2	Sequence 2, Appl
6	56	93.3	709	US-10-170-221-2	Sequence 2, Appl
7	56	93.3	711	US-10-169-297-8	Sequence 9, Appl
8	56	93.3	711	US-10-170-221-4	Sequence 4, Appl
9	56	93.3	711	US-10-341-834-202	Sequence 202, App
10	56	93.3	711	US-10-341-834-218	Sequence 218, App
11	56	93.3	711	US-10-440-464-69	Sequence 69, Appl
12	56	93.3	711	US-10-408-765A-2000	Sequence 2000, Ap
13	45	75.0	708	US-10-169-297-35	Sequence 35, Appl
14	45	75.0	708	US-10-169-297-8	Sequence 8, Appl
15	41	68.3	254	US-10-156-761-10952	Sequence 10952, A

16	40	66.7	123	US-10-437-663-151345	Sequence 151345,
17	40	66.7	295	US-10-437-663-113372	Sequence 113372,
18	40	66.7	328	US-09-978-295A-358	Sequence 358, App
19	40	66.7	328	US-09-978-697-358	Sequence 358, App
20	40	66.7	328	US-09-978-192A-358	Sequence 358, App
21	40	66.7	328	US-09-999-832A-358	Sequence 358, App
22	40	66.7	328	US-09-978-189-358	Sequence 358, App
23	40	66.7	328	US-09-978-608A-358	Sequence 358, App
24	40	66.7	328	US-09-978-585A-358	Sequence 358, App
25	40	66.7	328	US-09-978-191A-358	Sequence 358, App
26	40	66.7	328	US-09-978-403A-358	Sequence 358, App
27	40	66.7	328	US-09-978-664A-358	Sequence 358, App
28	40	66.7	328	US-09-999-832A-358	Sequence 358, App
29	40	66.7	328	US-09-981-915A-358	Sequence 358, App
30	40	66.7	328	US-09-978-824-358	Sequence 358, App
31	40	66.7	328	US-09-918-585A-358	Sequence 358, App
32	40	66.7	328	US-09-978-423A-358	Sequence 358, App
33	40	66.7	328	US-09-978-193A-358	Sequence 358, App
34	40	66.7	328	US-09-999-830A-358	Sequence 358, App
35	40	66.7	328	US-09-978-757A-358	Sequence 358, App
36	40	66.7	328	US-09-978-187B-358	Sequence 358, App
37	40	66.7	328	US-09-978-643A-358	Sequence 358, App
38	40	66.7	328	US-09-978-375A-358	Sequence 358, App
39	40	66.7	328	US-09-978-298A-358	Sequence 358, App
40	40	66.7	328	US-09-978-188A-358	Sequence 358, App
41	40	66.7	328	US-09-978-681A-358	Sequence 358, App
42	40	66.7	328	US-09-978-194A-358	Sequence 358, App
43	40	66.7	328	US-09-999-829A-358	Sequence 358, App
44	40	66.7	328	US-09-978-299A-358	Sequence 358, App
45	40	66.7	328	US-09-978-544A-358	Sequence 358, App

#### ALIGNMENTS

RESULT 1  
US-10-169-297-49  
Sequence 49, Application US/10169297  
Publication No. US2003017276A1  
GENERAL INFORMATION:  
APPLICANT: Tohdo, Naoki  
APPLICANT: Murata, Masashi  
APPLICANT: Enjoji, Takashi  
TITLE OF INVENTION: Preventives and Remedies for Chronic  
TITLE OF INVENTION: Hepatitis  
FILE REFERENCE: 3435.1000-000  
CURRENT APPLICATION NUMBER: US/10/169,297  
CURRENT FILING DATE: 2002-10-31  
PRIOR APPLICATION NUMBER: PCT/JP00/09393  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: JP 11/374087  
PRIOR FILING DATE: 1999-12-28  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 49  
LENGTH: 359  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-169-297-49  
Query Match 93.3%, Score 56, DB 14, Length 359;  
Best Local Similarity 91.7%; Pred. No. 0.19;  
Matches 11; Conservative 0; Mismatches 1; Indels 0;  
Gaps 0;  
Cy 1 ARRARYVMAVG 12  
Db 114 ARRARYVMAVG 125  
RESULT 2  
US-10-076-816-4  
Sequence 4, Application US/10076816  
Publication No. US20030056244A1

```

; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076,816
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,198
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-816-4

```

```

Query Match          93.3%; Score 56; DB 14; Length 690;
Best Local Similarity 91.7%; Pred. No. 0.35;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
QY      1  ARARVYMAVG 12
Db      339  ARARVYMAVG 350

```

```

RESULT 3
US-10-077-381-4
; Sequence 4, Application US/10077381
; Publication No. US2003074700A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Expression of Human Milk Proteins in
; FILE REFERENCE: 50665-8022.US00
; CURRENT APPLICATION NUMBER: US/10/077,381
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,199
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-077-381-4

```

```

Query Match          93.3%; Score 56; DB 14; Length 690;
Best Local Similarity 91.7%; Pred. No. 0.35;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
QY      1  ARARVYMAVG 12
Db      339  ARARVYMAVG 350

```

```

RESULT 4
US-10-639-835-4

```

```

; Sequence 4, Application US/10639835
; Publication No. US20040111766A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Expression of Human Milk Proteins in
; FILE REFERENCE: 50665-8022.US01
; CURRENT APPLICATION NUMBER: US/10/639,835
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: US 10/077,381
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,199
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-639-835-4

```

```

Query Match          93.3%; Score 56; DB 16; Length 690;
Best Local Similarity 91.7%; Pred. No. 0.35;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
QY      1  ARARVYMAVG 12
Db      339  ARARVYMAVG 350

```

```

RESULT 5
US-10-023-096-2
; Sequence 2, Application US/10023096
; Publication No. US20020160941A1
; GENERAL INFORMATION:
; APPLICANT: Krusel, Marian L.
; APPLICANT: Kurecki, Tomasz
; APPLICANT: Gollnick, Paul D.
; APPLICANT: Doyle, Darrell J.
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human
; CORRESPONDENCE ADDRESSES: 8
; ADDRESS: Jacobson, Price, Holman & Stern
; STREET: 400 Seventh St. N.W.
; CITY: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/023,096
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,586
; FILING DATE: 30-SEPT-1996
; APPLICATION NUMBER: US 08/238,445
; FILING DATE: 05-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409

```

REFERENCE/DOCKET NUMBER: 10505/P58185C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 93.3%; Score 56; DB 13; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.35;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
DB 343 ARRAVWCAVG 354

RESULT 6  
US-10-170-221-2  
Sequence 2, Application US/10170221  
Publication No. US20030192068A1  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
Strijker, Rein  
Heyneker, Herbert L.  
Platenburg, Gerald  
Lee, Sang He  
Pieper, Frank  
Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/170,221  
FILING DATE: 11-Jun-2002  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,798  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschuetz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-170-221-2

Query Match 93.3%; Score 56; DB 14; Length 709;  
Best Local Similarity 91.7%; Pred. No. 0.35;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
DB 358 ARRAVWCAVG 369

RESULT 7  
US-10-169-297-9  
Sequence 9, Application US/10169297  
Publication No. US20030171276A1  
GENERAL INFORMATION:  
APPLICANT: Tohdoh, Naoki  
APPLICANT: Murata, Masashi  
Enjoji, Takashi  
TITLE OF INVENTION: Preventives and Remedies for Chronic  
Hepatitis  
FILE REFERENCE: 3435.1000-000  
CURRENT APPLICATION NUMBER: US/10/169,297  
PRIORITY FILING DATE: 2002-10-31  
PRIOR APPLICATION NUMBER: PCT/JP00/09393  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: JP 11/374087  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 9  
LENGTH: 711  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-169-297-9

Query Match 93.3%; Score 56; DB 14; Length 711;  
Best Local Similarity 91.7%; Pred. No. 0.36;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
DB 360 ARRAVWCAVG 371

RESULT 8  
US-10-170-221-4  
Sequence 4, Application US/10170221  
Publication No. US20030192068A1  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
Strijker, Rein  
Heyneker, Herbert L.  
Platenburg, Gerald  
Lee, Sang He  
Pieper, Frank  
Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/170,221
  FILING DATE: 11-Jun-2002
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/476,798
  FILING DATE: 07-JUN-1995
  APPLICATION NUMBER: US 08/077,788
  FILING DATE: 15-JUN-1993
  APPLICATION NUMBER: US 07/895,956
  FILING DATE: 15-JUN-1992
  APPLICATION NUMBER: US 07/619,131
  FILING DATE: 27-NOV-1990
  APPLICATION NUMBER: US 07/444,745
  FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
  NAME: Liebeschultz, Joe O.
  REGISTRATION NUMBER: 37,505
  REFERENCE/DOCKET NUMBER: 16994-003125
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-543-9600
  TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 711 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-170-221-4

Query Match          93.3%; Score 56; DB 14; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVYVMAVG 12
Db 360 ARARVYVMAVG 371

RESULT 9
US-10-341-434-202
; Sequence 202, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 202
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-202

Query Match          93.3%; Score 56; DB 15; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVYVMAVG 12
Db 360 ARARVYVMAVG 371
```

```
RESULT 10
US-10-341-434-218
; Sequence 218, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 218
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-218

Query Match          93.3%; Score 56; DB 15; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVYVMAVG 12
Db 360 ARARVYVMAVG 371

RESULT 11
US-10-440-464-69
; Sequence 69, Application US/10440464
; Publication No. US20040018528A1
; GENERAL INFORMATION:
; APPLICANT: DERRINO, SAMUEL
; APPLICANT: O'FARRELL, ANNE-MARIE
; APPLICANT: MORIMOTO, AYESA
; APPLICANT: SMOLICH, BEVERLY
; APPLICANT: MANNING, WILLIAM
; APPLICANT: WALTER, SARAH
; APPLICANT: CHERRINGTON, JULIE
; TITLE OF INVENTION: NOVEL BIOMARKERS OF TYROSINE KINASE INHIBITOR EXPOSURE
; FILE REFERENCE: 038602/1592
; CURRENT APPLICATION NUMBER: US/10/440,464
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60/380,872
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/448,922
; PRIOR FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: 60/448,874
; PRIOR FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 69
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-440-464-69

Query Match          93.3%; Score 56; DB 15; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVYVMAVG 12
Db 360 ARARVYVMAVG 371
```

## RESULT 12

US-10-408-765A-2000  
; Sequence 2000, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Brian D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660098.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2000  
; LENGTH: 711  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-2000

## Query Match

93.3%; Score 56; DB 16; Length 711;  
Best Local Similarity 91.7%; Pred. No. 0.36;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ARARVWMAVG 12  
||| ||| |||

Db 360 ARARVWMAVG 371

## RESULT 13

US-10-169-297-35  
; Sequence 35, Application US/10169297  
; Publication No. US2003011276A1  
; GENERAL INFORMATION:  
; APPLICANT: Tohdoh, Naoki  
; APPLICANT: Murata, Masashi  
; APPLICANT: Enjoji, Takashi  
; TITLE OF INVENTION: Preventives and Remedies for Chronic  
; FILE REFERENCE: 3435.1000-000  
; CURRENT APPLICATION NUMBER: US/10/169,297  
; CURRENT FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: PCT/JP00/09393  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: JP 11/374087  
; PRIOR FILING DATE: 1999-12-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 234  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-10-169-297-35

## Query Match

75.0%; Score 45; DB 14; Length 234;  
Best Local Similarity 75.0%; Pred. No. 9.1;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARARVWMAVG 12  
||| ||| |||

Db 71 ARARVWMAVG 82

## RESULT 14

US-10-169-297-8  
; Sequence 8, Application US/10169297  
; Publication No. US2003011276A1  
; GENERAL INFORMATION:

## APPLICANT: Tohdoh, Naoki

; APPLICANT: Murata, Masashi  
; APPLICANT: Enjoji, Takashi  
; TITLE OF INVENTION: Preventives and Remedies for Chronic  
; FILE REFERENCE: 3435.1000-000  
; CURRENT APPLICATION NUMBER: US/10/169,297  
; CURRENT FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: PCT/JP00/09393  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: JP 11/374087  
; PRIOR FILING DATE: 1999-12-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 708  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-10-169-297-8

## Query Match

75.0%; Score 45; DB 14; Length 708;  
Best Local Similarity 75.0%; Pred. No. 25;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARARVWMAVG 12  
||| ||| |||

Db 359 ARARVWMAVG 370

## RESULT 15

US-10-156-761-10952  
; Sequence 10952, Application US/10156761  
; Publication No. US20030115018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 10952  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-10952

## Query Match

68.3%; Score 41; DB 14; Length 254;  
Best Local Similarity 80.0%; Pred. No. 46;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ARARVWMA 10  
||| ||| |||

Db 33 ARARVWMA 42

Search completed: September 1, 2004, 00:49:24  
Job time : 29.36 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2004, 00:03:45; Search time 6.96 Seconds  
(without alignments)

165,847 Million cell updates/sec

Title: US-09-508-095-22

Perfect score: 60

Sequence: 1 ABRARVMAAVG 12

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	93.3	711	1 TFFHL	lactotransferrin p
2	49	81.7	703	2 A4543	lactoferrin precu
3	45	75.0	708	1 TFFBL	lactotransferrin p
4	44	73.3	708	2 JCR23	lactoferrin - goat
5	43	71.7	491	2 T4684	p-hydroxybenzaldeh
6	40	71.7	707	1 A2843	lactoferrin precu
7	40	66.7	328	2 JEC375	carboxic anhydrase
8	39	66.7	449	2 H70526	probable cytochrom
9	39	65.0	285	1 IDECRP	replication initia
10	39	65.0	285	2 I64780	replication-associ
11	39	65.0	662	2 JCR906	sucrose 1P-fructos
12	38	63.3	176	2 F72863	hypothetical prote
13	38	63.3	307	2 AE3466	maltoase transport
14	38	63.3	509	2 T02860	beta-fructofuranos
15	38	63.3	986	2 A87590	hypothetical prote
16	37	61.7	108	2 F72507	hypothetical prote
17	37	61.7	338	2 T12816	hypothetical prote
18	37	61.7	379	2 AE3225	ABC transporter, m
19	37	61.7	397	2 A75503	hypothetical prote
20	37	61.7	419	1 S74500	zeaxanthin glucosyl
21	37	61.7	469	2 ABR700	hypothetical prote
22	37	61.7	477	2 D83617	probable amino aci
23	37	61.7	484	2 B97482	probable aminotran
24	37	61.7	651	2 A26581	beta-glucuronidase
25	36	60.0	77	2 AH1542	hypothetical prote
26	36	60.0	148	2 F86615	cytosine deaminase
27	36	60.0	155	2 E72007	cytosine deaminase
28	36	60.0	342	2 S76626	ABC-type transport
29	36	60.0	399	2 T30222	sensory protein Ki

30	36	60.0	447	2 E97460	hypothetical prote
31	36	60.0	497	2 T46876	proline uptake pro
32	36	60.0	501	1 P1WBP	IL protein - Europ
33	36	60.0	507	2 S36478	IL protein - human
34	36	60.0	512	2 E89978	high affinity prol
35	36	60.0	513	1 P1WLP	IL protein - deer
36	36	60.0	771	1 M2VZ4	ribonucleoside-dip
37	36	60.0	771	1 M2VZ5	ribonucleoside-dip
38	36	60.0	771	2 T28436	ribonucleoside-dip
39	36	60.0	771	2 B36843	ribonucleoside-dip
40	36	60.0	771	2 H72157	IL protein - vari
41	36	60.0	827	2 H83217	probable transcrip
42	36	60.0	1279	2 E64709	type IIS restricti
43	35	58.3	57	2 C41476	ribonucleoside-dip
44	35	58.3	84	2 D70967	hypothetical prote
45	35	58.3	133	2 A12834	conserved hypochet

## ALIGNMENTS

## RESULT 1

TFPHUL

lactotransferrin precursor [validated] - human

N.Alternate names: lactoferrin

C.Species: Homo sapiens (man)

C.Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C.Accession: G01394; S11228; A45401; S10324; S15853; S20541; S07160; A61169; A31000; S74

R.Chno, Y.

submitted to the EMBL Data Library, March 1994

A.Reference number: G06820

A.Accession: G01394

A>Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-711 <CHD>

A.Cross-references: EMBL:U07643; NID:9467236; PIDN:AAB60324.1; PID:9467237

R.Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A.Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A.Reference number: S11228; MUID:90384839; PMID:2402455

A.Accession: S11228

A.Molecule type: mRNA

A.Residues: 1-148, 'T', 150-422, 'C', 424-711 <REY>

A.Cross-references: EMBL:X53961; NID:934415; PIDN:CA93794.1; PID:934416

R.Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Pannella, T.

Mol. Endocrinol. 6, 1969-1991, 1992

A.Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A.Reference number: A45401; MUID:93125571; PMID:1480183

A.Accession: A45401

A.Molecule type: DNA

A.Residues: 1-15 <TEN>

A.Cross-references: GB:S52659; NID:9263311; PIDN:AAB24877.1; PID:9263312

A.Experimental source: placenta

A.Note: sequence extracted from NCBI backbone (NCBI:P1.22202)

R.Powell, M.J.; Ogden, J.E

Nucleic Acids Res. 18, 4013, 1990

A.Title: Nucleotide sequence of human lactoferrin cDNA.

A.Reference number: S10324; MUID:90326549; PMID:2374734

A.Accession: S10324

A.Molecule type: mRNA

A.Residues: 3-711 <POW>

A.Cross-references: EMBL:X52941; NID:934411; PIDN:CA937116.1; PID:934412

R.Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A>Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A.Reference number: S15853; MUID:91264786; PMID:2049066

A.Accession: S15853

A.Molecule type: mRNA

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A.Residues: 20-31 <ST1>

A.Accession: S20841

A.Molecule type: protein

A.Residues: 20-28, 'X', 30-31 <ST2>

R/Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 A/Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mR  
 A/Reference number: S07160; PMID:88001031; PMID:3477300  
 A/Accession: S07160  
 A/Molecule type: mRNA  
 A/Residues: 436-487, 'A', 489-711 <RAD>  
 A/Cross-references: EMBL:M8642; NID:9186815; PIDN:AAA8665.1; PID:938655  
 R/Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 A/Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocy  
 A/Reference number: A61169; PMID:91235214; PMID:5674448  
 A/Accession: A61169  
 A/Status: not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 3-701, 'SMKPVN' <PAN>  
 A/Experimental source: normal breast tissue  
 R/Netz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A/Title: Human lactotransferrin: amino acid sequence and structural comparisons with oth  
 A/Reference number: A31000; PMID:85076667; PMID:6510420  
 A/Accession: A31000  
 A/Molecule type: protein  
 A/Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-365, 'Q', 387-391, 'W', 393-4  
 A/Note: this is the final paper in a series  
 R/Houen, G.; Hoegdall, E.V.; Barkholt, V.; Nørskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A/Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affini  
 A/Reference number: S74119; PMID:97054624; PMID:8898921  
 A/Accession: S74119  
 A/Molecule type: protein  
 A/Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A/Experimental source: neutrophil granulocytes  
 C/Genetics:  
 A/Gene: GDB:LTf  
 A/Cross-references: GDB:119368; OMIM:150210  
 A/Map position: 3q21-3q23  
 C/Superfamily: transferrin; transferrin repeat homology  
 C/Keywords: duplication; glycoprotein; iron binding; milk  
 F/1-19/Domain: signal sequence #status predicted <SIG>  
 F/20-711/Product: lactotransferrin #status experimental <NAT>  
 F/21-356/Domain: transferrin repeat homology <TRH1>  
 F/360-699/Domain: transferrin repeat homology <TRH2>  
 F/39-66-39-56-133-218-177-193-190-201-251-265-503-697-595-609/Disulfide bonds: #status e  
 F/557-498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F/368-400-378-351, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stat  
 Query Match 93.3%; Score 56; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.031;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 2  
 A/Accession: A45543  
 A/Title: Lactoferrin precursor - pig  
 C/Species: Sus scrofa domestica (domestic pig)  
 C/Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 04-Mar-2000  
 A/Accession: A45543; S24173  
 R/Alexander, L.J.; Levine, W.B.; Teng, C.T.; Beattie, C.W.  
 Anim. Genet. 23, 251-256, 1992  
 A/Title: Cloning and sequencing of the porcine lactoferrin cDNA.  
 A/Reference number: A45543; PMID:92367939; PMID:1503255  
 A/Accession: A45543  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-703 <ALB>  
 A/Experimental source: mammary gland  
 A/Note: sequence extracted from NCBI backbone (NCBIN:111151, NCBI:111153)  
 R/Lydon, J.P.; O'Malley, B.R.; Saucedo, O.; Lee, T.; Headon, D.R.; Connely, O.M.

Biochim. Biophys. Acta 1132, 97-99, 1992  
 A/Title: Nucleotide and primary amino acid sequence of porcine lactoferrin.  
 A/Reference number: S24173; PMID:92379101; PMID:1511016  
 A/Accession: S24173  
 A/Molecule type: mRNA  
 A/Residues: 1-11, 'W', 13-50, 'I', 52-84, 'G', 86-120, 'L', 121-130, 'I', 132-282, 'S', 284-571, 'Q',  
 A/Cross-references: EMBL:M92089; NID:9164613; PIDN:AAA31102.1; PID:9164614  
 A/Experimental source: mammary gland  
 C/Superfamily: transferrin; transferrin repeat homology  
 C/Keywords: duplication; glycoprotein; iron; iron binding; metalloprotein; milk  
 F/1-19/Domain: signal sequence #status predicted <SIG>  
 F/20-703/Product: lactoferrin #status predicted <NAT>  
 F/20-350/Domain: transferrin repeat homology <TRH1>  
 F/36-48/Region: antimicrobial  
 F/354-691/Domain: transferrin repeat homology <TRH2>  
 F/28-62-38-53-129-212-171-187-184-195-245-259-362-394-372-385-419-698-439-661-471-546-49  
 F/77-107-206-267/Binding site: iron (Asp, Tyr, Tyr, His) #status predicted  
 F/103/Binding site: carbonate (Arg) #status predicted  
 F/409, 447, 540, 609/Binding site: iron (Asp, Tyr, Tyr, His) #status predicted  
 F/477/Binding site: carbonate (Arg) #status predicted  
 F/490/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 81.7%; Score 49; DB 2; Length 703;  
 Best Local Similarity 75.0%; Pred. No. 0.56;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 3  
 A/Accession: A45543  
 A/Title: Lactoferrin precursor - bovine  
 C/Species: Bos primigenius taurus (cattle)  
 C/Date: 31-Mar-1992 #sequence\_revision 21-Nov-1997 #text\_change 11-May-2000  
 A/Accession: A45543; S14674; S14110; S18517; U05555; S13097; S18518; S13881; E01048; S21  
 R/Tsang, T.C.; Burns, D.K.; Wang, F.; Pan, Y.  
 FASEB J. 6, 233, 1991  
 A/Title: Cloning of a 80-kD advanced glycosylation end product (AGE) binding protein fro  
 A/Reference number: A45919  
 A/Accession: A45919  
 A/Status: translated from GB/EMBL/DDBT  
 A/Molecule type: mRNA  
 A/Residues: 1-708 <ISA>  
 A/Cross-references: GB:D08604; NID:9163269; PIDN:AAA30609.1; PID:9163270  
 R/Pierce, A.  
 submitted to the EMBL Data Library, November 1990  
 A/Reference number: S14674  
 A/Accession: S14674  
 A/Molecule type: mRNA  
 A/Residues: 1-144, 'V', 146-163, 'P', 166-339, 'A', 341-438, 'Y', 440-513, 'R', 515-708 <PI>  
 A/Cross-references: EMBL:X57084; NID:9505; PIDN:CAA40366.1; PID:9506  
 R/Pierce, A.; Colavizza, D.; Benalissa, M.; Maes, P.; Tataru, A.; Montreuil, J.; Spik, G.  
 Eur. J. Biochem. 196, 177-184, 1991  
 A/Title: Molecular cloning and sequence analysis of bovine lactotransferrin.  
 A/Reference number: S14110; PMID:91160550; PMID:2001696  
 A/Accession: S14110  
 A/Molecule type: mRNA  
 A/Residues: 3-144, 'V', 146-339, 'A', 341-438, 'Y', 440-513, 'R', 515-708 <PI>  
 A/Cross-references: EMBL:X57084  
 A/Accession: S18517  
 A/Molecule type: protein  
 A/Residues: 20-35, 82-114, 148-163, 'P', 166-178, 'V', 'P', 183-190, 205-212, 230-239, 304-339, 59  
 R/Goodman, R.R.; Schandbacher, F.L.  
 Biochem. Biophys. Res. Commun. 180, 75-84, 1991  
 A/Title: Bovine lactoferrin mRNA: sequence, analysis, and expression in the mammary glan  
 A/Reference number: U05955; PMID:92028986; PMID:1718281  
 A/Accession: U05955  
 A/Molecule type: mRNA  
 A/Residues: 1-65, 'PG', 68-296, 'S', 298-339, 'A', 341-708 <GOO>  
 A/Cross-references: GB:M63502



A>Note: the authors translated the codon CCG for residue 66 as Arg and TCT for residue 2  
R:Mead, P.E.; Tweedle, J.W.  
Nucleic Acids Res. 18, 7167, 1990  
A>Title: cDNA and protein sequence of bovine lactoferrin.  
A:Reference number: S13097; MUID:91086328; PMID:2263492  
A:Accession: S13097  
A:Molecule type: mRNA  
A:Residues: 28-33, 'DS', 36-38, 'P', 40-708 <MEA>  
A:Cross-references: EMBL:X54801  
A:Accession: S18518  
A:Molecule type: protein  
A:Residues: 20-47,55-66;132-139;256-277;278,305-332;343-351,361-363;586,587-589;598-619  
R:Read, P.B.  
submitted to the EMBL Data Library, October 1990  
A:Reference number: S13881  
A:Accession: S13881  
A:Molecule type: mRNA  
A:Residues: 26-38, 'P', 40-86, 'C', 88-708 <ME3>  
A:Cross-references: EMBL:X54801  
R:Rejman, J.U.; Hegarty, H.M.; Hurley, W.L.  
Comp. Biochem. Physiol. B 93, 929-934, 1989  
A>Title: Purification and characterization of bovine lactoferrin from secretions of the  
A:Reference number: PLO148; MUID:90031466; PMID:2805645  
A:Accession: PLO148  
A:Molecule type: protein  
A:Residues: 20-27, 'X', 29-37, 'X', 39-54, 'X', 56-59 <REJ>  
R:bellamy, W.; Takase, M.; Yamauchi, K.; Wakabayashi, H.; Kawase, K.; Tomita, M.  
Biochim. Biophys. Acta 1121, 130-136, 1992  
A>Title: Identification of the bactericidal domain of lactoferrin.  
A:Reference number: S21756; MUID:92287941; PMID:159934  
A:Accession: S21756  
A:Molecule type: protein  
A:Residues: 36-60 <EBL>  
R:Shimazaki, K.; Tanaka, T.; Kon, H.; Oota, K.; Kawaguchi, A.; Maki, Y.; Sato, T.  
J. Dairy Sci. 76, 946-955, 1993  
A>Title: Separation and characterization of the C-terminal half molecule of bovine lacto  
A:Reference number: A56659; MUID:93253156; PMID:848645  
A:Accession: A56659  
A:Molecule type: protein  
A:Residues: 20-25;302-308;359-366, 'X', 368-376, 'X', 378 <SHI>  
C:Superfamily: transferrin; transferrin repeat homology  
C:Keywords: duplication; glycoprotein; iron; iron binding; metalloprotein; milk  
F:1-19/Domains: signal sequence #status predicted <SIG>  
F:20-708/Product: lactoferrin repeat homology <TRH>  
F:20-355/Domains: transferrin repeat homology <TRH>  
F:356-60/Region: antimicrobial  
F:359-696/Domains: transferrin repeat homology <TRH>  
F:28-64,134-217,176-192,179-200,189-202,250-264,367-399,377-390,424-703,444-666,476-551,  
F:38-55/Distulfide bonds: #status predicted  
F:79,111,211,272/Binding site: iron (Asp, Tyr, Tyr, His) #status experimental  
F:140/Binding site: carbohydrate (Arg) #status experimental  
F:252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:414,452,545,614/Binding site: iron (Asp, Tyr, Tyr, His) #status experimental  
F:482/Binding site: carbonate (Arg) #status experimental

Query Match 75.0%; Score 45; DB 1; Length 708;  
Best Local Similarity 75.0%; Pred. No. 3;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
DB 359 ARCTRVWCAVG 370

RESULT 4  
JC2323  
lactoferrin - goat  
C:Species: Capra aegagrus hircus (domestic goat)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
C:Accession: JC2323  
R:Le Provost, F.; Nocard, M.; Guerin, G.; Martin, P.  
Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
A>Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus

A:Reference number: JC2323; MUID:94380047; PMID:8093048  
A:Accession: JC2323  
A:Molecule type: mRNA  
A:Residues: 1-708 <LEP>  
C:Superfamily: transferrin; transferrin repeat homology  
C:Keywords: duplication; glycoprotein  
F:359-696/Domains: transferrin repeat homology <TRH>  
F:252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.3%; Score 44; DB 2; Length 708;  
Best Local Similarity 75.0%; Pred. No. 4.5;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
DB 359 ARCTRVWCAVG 370

RESULT 5  
T6684  
p-hydroxybenzaldehyde dehydrogenase [imported] - Pseudomonas putida plasmid pRA4000  
C:Species: Pseudomonas putida  
C>Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000  
C:Accession: T6684  
R:Cronin, C.N.; Kim, J.H.; Fuller, J.; Zhang, X.P.; McIntire, W.S.  
DNA Seq. 10, 7-17, 1999  
A>Title: Organization and sequences of p-hydroxybenzaldehyde dehydrogenase and other pla  
869.  
A:Reference number: Z23132; MUID:20029269; PMID:10565539  
A:Accession: T46684  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-491 <CRO>  
A:Cross-references: EMBL:U96338; NID:94808503; PIDN:AAA75634.2; PID:G4808506  
A:Experimental source: NCIMB 9866  
C:Genetics:  
A:Gene: pcba  
A:Genome: plasmid pRA4000  
C:Function:  
A:Description: catalyzes oxidation of p-hydroxybenzaldehyde to p-hydroxybenzoic acid  
A>Note: cofactor NADP.  
C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 71.7%; Score 43; DB 2; Length 491;  
Best Local Similarity 66.7%; Pred. No. 4.8;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
DB 58 ARQAQVMAAG 69

RESULT 6  
A28438  
lactoferrin precursor - mouse  
N:Alternate names: lactotransferrin  
C:Species: Mus musculus (house mouse)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A28438, A41205  
R:Pentecost, B.T.; Teng, C.T.  
J. Biol. Chem. 262, 10134-10139, 1987  
A>Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secre  
A:Reference number: A92596; MUID:87280033; PMID:3611056  
A:Accession: A28438  
A:Molecule type: mRNA  
A:Residues: 3-707 <PEN>  
A:Cross-references: EMBL:J03298  
R:Lin, Y.; Teng, C.T.  
J. Biol. Chem. 266, 21880-21885, 1991  
A>Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
A:Reference number: A41205; MUID:92042099; PMID:1939212  
A:Accession: A41205  
A:Molecule type: DNA

A:Residues: 1-15 <LIU>  
A:Cross-references: GB:M74778  
C:Superfamily: transferrin; transferrin repeat homology  
C:Keywords: duplication; glycoprotein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-707/Product: lactoferrin #status predicted <MNT>  
F:358-695/Domain: transferrin repeat homology <TRH2>  
F:494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.7%; Score 43; DB 1; Length 707;  
Best Local Similarity 66.7%; Pred. No. 6.8;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARARVMAAVG 12  
DB 358 ASKARVMAVG 369

RESULT 7  
JE0375  
Carbonic anhydrase-related protein - human

C:Species: Homo sapiens (man)  
C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: JE0375  
R:Beltingham, J.; Gregory-Evans, K.; Y. Gregory-Evans, C.  
Biochem. Biophys. Res. Commun. 253, 364-367, 1998  
A:Title: Sequence and tissue expression of a novel human carbonic anhydrase-related protein  
A:Reference number: JE0375; MIM:19097242; PMID:9878543  
A:Accession: JE0375  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-328 <BEL>  
A:Cross-references: GB:A067662; NID:93283385; PIDN:AA09689.1; PID:93283386  
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology  
F:45-303/Domain: carbonic anhydrase homology <CAH>

Query Match 66.7%; Score 40; DB 2; Length 328;  
Best Local Similarity 66.7%; Pred. No. 11;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARARVMAAVG 12  
DB 9 APRALVMAVG 20

## RESULT 8

H70526  
Probable cytochrome P450 RV0327C - Mycobacterium tuberculosis (strain H37RV)  
N:Contains: oxidoreductase (BC 1.-.-.-)  
C:Species: Mycobacterium tuberculosis  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 23-Dec-2002  
C:Accession: H70526  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Rajandream, A.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MIM:98295987; PMID:9634230  
A:Accession: H70526  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-449 <COL>  
A:Cross-references: GB:296800; GB:AL123456; NID:93261800; PIDN:CAM09576.1; PID:92193948  
A:Experimental source: strain H37RV  
C:Genetics:

A:Gene: RV0327C  
C:Superfamily: Bacillus cytochrome P450 CYP106; cytochrome P450 homology  
C:Keywords: chromoprotein, heme, iron, metalloprotein; oxidoreductase  
F:259-405/Domain: cytochrome P450 homology <P45>  
F:383/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 66.7%; Score 40; DB 2; Length 449;

Best Local Similarity 70.0%; Pred. No. 15;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RARVMAAVG 12  
DB 285 QSOVMAAVG 294

## RESULT 9

IDECRP  
replication initiation protein - Escherichia coli plasmids  
C:Species: Escherichia coli  
C:Date: 18-Dec-1981 #sequence\_revision 17-Dec-1982 #text\_change 16-Jul-1999  
C:Accession: A03602; B03602; S01773; S05591; A48662; C28378; I41109  
R:Rosen, J.; Ryder, T.; Inokuchi, H.; Ohtsubo, H.; Ohtsubo, E.  
Mol. Gen. Genet. 179, 527-537, 1980  
A:Title: Genes and sites involved in replication and incompatibility of an R100 plasmid  
A:Reference number: A93119; MIM:81074309; PMID:7003300  
A:Accession: A03602  
A:Molecule type: DNA  
A:Residues: 1-285 <ROI>  
A:Cross-references: GB:J01762; GB:J01761; GB:J01767; GB:J01768; NID:9151740; PIDN:AAA322  
A:Experimental source: Plasmid R100  
R:Rosen, J.; Ryder, T.; Ohtsubo, H.; Ohtsubo, E.  
Nature 290, 794-797, 1981  
A:Title: Role of RNA transcripts in replication incompatibility and copy number control  
A:Reference number: A93253; MIM:81173118; PMID:6163994  
A:Accession: B03602

A:Molecule type: DNA  
A:Residues: 1-77 <ROS>  
A:Experimental source: plasmid R1  
R:Dong, X.; Womble, D.D.; Rownd, R.H.  
J. Mol. Biol. 202, 495-509, 1988  
A:Title: In-vivo studies on the cis-acting replication initiator protein of IncFII plasmid  
A:Reference number: S01773; MIM:89011975; PMID:3050127  
A:Accession: S01773  
A:Molecule type: DNA  
A:Residues: 1-285 <CON>  
A:Cross-references: EMBL:X12776; NID:942707; PIDN:CAA31263.1; PID:9581213  
A:Experimental source: Plasmid NRI  
R:Masai, H.; Arai, K.I.  
Nucleic Acids Res. 16, 6493-6514, 1988  
A:Title: Repa protein- and oriR-dependent initiation of R1 plasmid replication: identification  
A:Reference number: S05591; MIM:88289416; PMID:3041379  
A:Accession: S05591  
A:Molecule type: DNA  
A:Status: translation not shown

A:Molecule type: DNA  
A:Residues: 270-285 <MAS>  
A:Cross-references: EMBL:X12587; NID:941107; PIDN:CAA31100.1; PID:941108  
A:Experimental source: Plasmid R1  
R:Ulang, T.; Min, Y.N.; Liu, W.; Womble, D.D.; Rownd, R.H.  
J. Bacteriol. 175, 5350-5358, 1993  
A:Title: Insertion and deletion mutations in the repA region of the IncFII plasmid NRI  
A:Reference number: A48662; MIM:93374828; PMID:8396115  
A:Accession: A48662  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 252-285 <JIA>  
A:Experimental source: Plasmid NRI  
R:Dong, X.; Womble, D.D.; Rownd, R.H.  
J. Bacteriol. 169, 5353-5363, 1987  
A:Title: Transcriptional pausing in a region important for plasmid NRI replication control  
A:Reference number: A28378; MIM:88058738; PMID:2445727  
A:Accession: C28378  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-50 <DO2>  
R:Womble, D.D.; Sampathkumar, P.; Easton, A.M.; Luckew, V.A.; Rownd, R.H.  
J. Mol. Biol. 181, 395-410, 1995  
A:Title: Transcription of the replication control region of the IncFII R-plasmid NRI in  
A:Reference number: I41109; MIM:85160860; PMID:2580099  
A:Accession: I41109  
A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA  
A:Residues: 1-285 <RES>  
A:Cross-references: EMBL:X02302; NID:g42132; PIDD:CA26168.1; PID:g581144  
C:Genetics:  
A:Gene: repA1  
A:Genome: plasmid  
A:Start codon: GTG  
C:Superfamily: rep1 protein  
C:Keywords: plasmid copy control

Query Match 65.0%; Score 39; DB 1; Length 285;  
Best Local Similarity 87.5%; Pred. No. 15;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRARVW 8  
|||:||||  
Db 168 ARRSRVW 175

## RESULT 10

164780  
replication-associated protein A1 - Escherichia coli  
C:Species: Escherichia coli  
C>Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 26-Aug-1999  
C:Accession: 164780; 14110  
R:Ohtsubo, H.; Ryder, T.B.; Maeda, Y.; Armstrong, K.A.; Ohtsubo, E.  
Adv. Biophys. 21, 115-133, 1986  
A:Title: DNA replication of the resistance plasmid R100 and its control.  
A:Reference number: 151821; PMID:86319522; PMID:3019092  
A:Accession: 164780  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-285 <RES>  
A:Cross-references: GB:M26840; NID:g151770; PIDD:AA26067.1; PID:g151773  
R:Womble, D.D.; Samathkumar, P.; Easton, A.M.; Luckow, V.A.; Rownd, R.H.  
J. Mol. Biol. 181, 395-410, 1985  
A:Title: Transcription of the replication control region of the IncFII R-plasmid NR1 in  
A:Reference number: 141106; PMID:8510860; PMID:2580099  
A:Accession: 14110  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 'M', 166-285 <RE2>  
A:Cross-references: EMBL:X02302; NID:g42132; PIDD:CA26169.1; PID:g581145  
C:Superfamily: rep1 protein

Query Match 65.0%; Score 39; DB 2; Length 285;  
Best Local Similarity 87.5%; Pred. No. 15;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRARVW 8  
|||:||||  
Db 168 ARRSRVW 175

## RESULT 11

JC7906  
sucrose 1F-fructosyltransferase (EC 2.4.1.99) - wheat  
N:Alternate names: 1-SST; fructosyltransferase 2; Wft2 protein  
C:Species: Triticum aestivum (common wheat)  
C>Date: 03-Feb-2003 #sequence\_revision 03-Feb-2003 #text\_change 14-Apr-2003  
C:Accession: JC7906  
R:Kawakami, A.; Yoshida, M.  
Biosci. Biotechnol. Biochem. 66, 2297-2305, 2002  
A:Title: Molecular characterization of sucrose 1-fructosyltransferase and sucrose  
A:Reference number: JC7905; PMID:12334550; PMID:12506964  
A:Accession: JC7906  
A:Molecule type: mRNA  
A:Residues: 1-662 <KAW>  
A:Cross-references: DDBJ:AB02988  
C:Comment: This enzyme, which is a vacuole-type fructosyltransferase and a fructan-biosy  
fers a fructosyl moiety from one sucrose to another, resulting in the formation of trisac  
C:Genetics:  
A:Gene: wft2

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 65.0%; Score 39; DB 2; Length 662;  
Best Local Similarity 63.6%; Pred. No. 34;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRARVMAVG 12  
:::|||||  
Db 402 KKRVMAYVG 412

## RESULT 12

P72563  
hypothetical protein APE1792 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: F72563  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takat  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, C.; I  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; PMID:99310339; PMID:10382966  
A:Accession: F72563  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-176 <KAW>  
A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDD:BA80795.1; PID:g5105482  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1792  
C:Superfamily: conserved hypothetical protein M1568

Query Match 63.3%; Score 38; DB 2; Length 176;  
Best Local Similarity 70.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 RARVMAVG 12  
|||:||||  
Db 72 RPRVWVG 81

## RESULT 13

A83466  
maltose transport system permease protein malP [imported] - Brucella melitensis (strain  
C:Species: Brucella melitensis  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
C:Accession: A83466  
R:DelVecchio, V.G.; Kapural, V.; Redkar, R.V.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
J.; Mazur, M.; Goldsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letesi  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A:Reference number: AD3252; PMID:1175688  
A:Accession: A83466  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-307 <KUR>  
A:Cross-references: GB:AF006917; PIDD:AL52896.1; PID:g17983741; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BME11715  
A:Map position: 1  
C:Superfamily: inner membrane protein ugpA

Query Match 63.3%; Score 38; DB 2; Length 307;  
Best Local Similarity 50.0%; Pred. No. 25;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARRARVMAVG 12  
|||:||||  
Db 251 ARRTLWTVG 262

## RESULT 14

T02260  
 beta-fructofuranosidase (EC 3.2.1.26) - maize (fragment)  
 N:Alternate names: invertase  
 C:Species: Zea mays (maize)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 22-Jun-1999  
 C:Accession: T02260  
 R:Xu, J.; Almlira, E.C.; Avigne, W.T.; McCarthy, D.R.; Koch, K.E.  
 Submitted to the EMBL Data Library, July 1995  
 A:Description: Molecular characterization and differential expression of a invertase gen  
 A:Reference number: Z14645  
 A:Accession: T02260  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-509 <RLJ>  
 A:Cross-references: EMBL:U31451; NID:g951165; PIDD:AAA74584.1; PID:g951166  
 A:Experimental source: strain merit; root tip  
 C:Function:  
 A:Description: catalyzes hydrolyzation of terminal non-reducing beta-D-fructofuranoside  
 C:Superfamily: beta-fructofuranosidase  
 C:Keywords: glycosidase; hydrolase

Query Match 63.3%; Score 38; DB 2; Length 509;  
 Best Local Similarity 58.3%; Pred. No. 40;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRRVWMAVG 12  
 | : | | : | | |  
 Db 248 AKRRRVLMGWS 259

RESULT 15  
 A87590  
 hypothetical protein CC2753 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: A87590  
 R:Merlan, W.C.; Peidoblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: A87590  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-986 <STO>  
 A:Cross-references: GB:AE005673; NID:g13424347; PIDD:AKX24717.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC2753

Query Match 63.3%; Score 38; DB 2; Length 986;  
 Best Local Similarity 87.5%; Pred. No. 74;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 RARVWAA 10  
 | | | | | | | |  
 Db 810 RARVWAS 817

Search completed: September 1, 2004, 00:17:08  
 Job time : 8.96 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2004, 23:54:44 ; Search time 4.08 Seconds

(without alignments)  
153.148 Million cell updates/sec

Title: US-09-508-095-22

Perfect score: 60

Sequence: 1 ARARVMAAVG 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	93.3	708	1 TRFL_BUBBU	O77698 bubalis bub
2	56	93.3	711	1 TRFL_HUMAN	P02788 homo sapien
3	51	85.0	695	1 TRFL_HORSE	O77811 equus caball
4	49	81.7	704	1 TRFL_PIG	P14632 sus scrofa
5	48	80.0	708	1 TRFL_CAMDR	Q91600 camelus dro
6	45	75.0	708	1 TRFL_BOVIN	P24627 bos taurus
7	44	73.3	708	1 TRFL_CAPI	Q29477 capra hircu
8	43	71.7	707	1 TRFL_MOUSE	P08071 mus musculu
9	40	66.7	449	1 CABH_HUMAN	O75493 homo sapien
10	40	66.7	449	1 CL3A_MYCTU	O08447 mycobacteri
11	39	65.0	285	1 REP2_ECOLI	P03066 escherichia
12	38	63.3	221	1 YH92_AERPE	Q9Y003 aeropyrum p
13	37	61.7	716	1 GID5_BIFLO	O89514 bitidobacte
14	37	61.7	368	1 GALT_HUMAN	O60755 homo sapien
15	37	61.7	651	1 BGLR_HUMAN	P08236 homo sapien
16	36	60.0	328	1 CABH_MOUSE	O70354 mus musculu
17	36	60.0	328	1 CABH_SHEEP	Q95003 ovis aries
18	36	60.0	501	1 VLI_PAPVE	P11326 european el
19	36	60.0	507	1 VLI_HPV15	Q05137 human papill
20	36	60.0	513	1 VLI_PAPVD	P03104 deer papill
21	36	60.0	771	1 RIRI_VACCC	P20503 vaccinia vi
22	36	60.0	771	1 RIRI_VACCV	P12848 vaccinia vi
23	36	60.0	771	1 RIRI_VAVV	P32384 variola vir
24	36	60.0	895	1 GNDS_RAT	Q03386 rattus norv
25	36	60.0	914	1 GNDS_HUMAN	Q12967 homo sapien
26	35	58.3	228	1 RS3_THETH	P08372 thermus the
27	35	58.3	324	1 GSHB_GLOVI	Q7M144 gloebacter
28	35	58.3	338	1 GPDA_SHEON	O8EKD9 shewanella
29	35	58.3	339	1 GPDA_YEPE	O8EKD9 yersinia pe
30	35	58.3	340	1 GPDA_SEEMA	O8K1M1 serratia ma
31	35	58.3	495	1 VLI_BPV1	P03101 bovine papl
32	35	58.3	497	1 VLI_BPV2	P06458 bovine papl
33	35	58.3	499	1 VLI_HPV33	P06416 human papil

## ALIGNMENTS

```

RESULT 1
TRFL_BUBBU          STANDARD;          PRT;          708 AA.
AC O77698;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin).
GN LTF.
OS Bubalis bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE FROM N.A.
RA Paramasivam M., Thattaiyath B.D., Kumar A., Srinivasan A.,
RA Singh T.P.;
RT "CDNA sequence of Buffalo lactoferrin."
RU Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=20003130; PubMed=10531476;
RA Karthikeyan S., Paramasivam M., Yadav S., Srinivasan A., Singh T.P.;
RT "Structure of buffalo lactoferrin at 2.5-A resolution using crystals
RT grown at 303 K shows different orientations of the N and C lobes."
RU Acta Crystallogr. D 55:1805-1813(1999).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: Composed of two homologous domains.
CC -!- SIMILARITY: Belongs to the transferrin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC -----
CC EMBL: AJ005203; CAA06441.1; -.
CC PDB: 1CE2; 19-MAR-99.
CC PDB: 1B1V; 13-JAN-99.
CC InterPro: IPR001156; Transferrin.
CC Pfam: PF00405; Transferrin.
CC PRINTS: PR00442; TRANSFERRIN.
CC SMART: SM00094; TR_FER; 2.
CC PROSITE: PS00205; TRANSFERRIN_1; 2.
CC PROSITE: PS00206; TRANSFERRIN_2; 2.
CC PROSITE: PS00207; TRANSFERRIN_3; 2.
CC Transport: iron transport; Glycoprotein; Metal-binding; Repeat;
CC Signal: 3d-structure.
CC PT SIGNAL 1 19

```

FT	CHAIN	20	708	
FT	DISULFID	28	64	
FT	DISULFID	38	55	
FT	DISULFID	134	217	
FT	DISULFID	176	192	
FT	DISULFID	179	202	
FT	DISULFID	189	200	
FT	DISULFID	250	264	
FT	DISULFID	367	399	
FT	DISULFID	377	390	
FT	DISULFID	424	703	
FT	DISULFID	444	666	
FT	DISULFID	476	551	
FT	DISULFID	500	694	
FT	DISULFID	510	524	
FT	DISULFID	521	534	
FT	DISULFID	592	534	
FT	DISULFID	644	606	
FT	METAL	79	79	IRON 1.
FT	METAL	111	111	IRON 1.
FT	METAL	211	211	IRON 1.
FT	METAL	272	272	IRON 1.
FT	METAL	414	414	IRON 2.
FT	METAL	452	452	IRON 2.
FT	METAL	545	545	IRON 2.
FT	METAL	614	614	IRON 2.
FT	BINDING	136	136	CARBONATE 1.
FT	BINDING	140	140	CARBONATE 1 (VIA AMIDE NITROGEN).
FT	BINDING	142	142	CARBONATE 1 (VIA AMIDE NITROGEN).
FT	BINDING	143	143	CARBONATE 1 (VIA AMIDE NITROGEN).
FT	BINDING	478	478	CARBONATE 2.
FT	BINDING	482	482	CARBONATE 2 (VIA AMIDE NITROGEN).
FT	BINDING	484	484	CARBONATE 2 (VIA AMIDE NITROGEN).
FT	BINDING	485	485	CARBONATE 2 (VIA AMIDE NITROGEN).
FT	CARBOHYD	252	252	N-LINKED (GLCMAC. . .) (POTENTIAL).
FT	CARBOHYD	300	300	N-LINKED (GLCMAC. . .) (POTENTIAL).
FT	CARBOHYD	495	495	N-LINKED (GLCMAC. . .) (POTENTIAL).
FT	CARBOHYD	564	564	N-LINKED (GLCMAC. . .) (POTENTIAL).
FT	TURN	21	22	
FT	STRAND	25	29	
FT	HELIX	32	46	
FT	TURN	47	48	
FT	STRAND	53	57	
FT	HELIX	61	69	
FT	TURN	70	71	
FT	STRAND	75	78	
FT	HELIX	80	87	
FT	TURN	89	91	
FT	STRAND	93	102	
FT	STRAND	107	108	
FT	STRAND	110	118	
FT	TURN	119	120	
FT	HELIX	125	127	
FT	TURN	129	130	
FT	STRAND	132	135	
FT	TURN	136	139	
FT	TURN	141	144	
FT	HELIX	145	150	
FT	TURN	151	151	
FT	HELIX	152	155	
FT	TURN	159	161	
FT	HELIX	164	169	
FT	TURN	170	171	
FT	STRAND	174	176	
FT	TURN	178	179	
FT	TURN	182	184	
FT	HELIX	186	189	
FT	TURN	190	191	
FT	HELIX	196	198	
FT	TURN	199	200	
FT	TURN	204	205	
FT	TURN	207	208	
FT	HELIX	211	219	

FT	TURN	220	221	
FT	STRAND	225	229	
FT	TURN	230	231	
FT	HELIX	232	236	
FT	TURN	240	243	
FT	STRAND	244	245	
FT	TURN	246	249	
FT	STRAND	251	252	
FT	STRAND	255	257	
FT	HELIX	258	260	
FT	TURN	261	263	
FT	STRAND	267	270	
FT	STRAND	273	277	
FT	HELIX	283	297	
FT	TURN	299	300	
FT	TURN	307	308	
FT	TURN	312	313	
FT	TURN	321	322	
FT	STRAND	325	328	
FT	TURN	331	332	
FT	HELIX	335	351	
FT	STRAND	354	362	
FT	STRAND	364	369	
FT	HELIX	371	384	
FT	TURN	385	385	
FT	STRAND	388	393	
FT	HELIX	396	404	
FT	TURN	405	406	
FT	STRAND	410	413	
FT	HELIX	415	422	
FT	TURN	423	425	
FT	STRAND	426	434	
FT	TURN	438	441	
FT	TURN	444	446	
FT	STRAND	452	459	
FT	TURN	460	461	
FT	HELIX	463	464	
FT	TURN	467	469	
FT	STRAND	471	472	
FT	TURN	474	477	
FT	TURN	480	481	
FT	HELIX	483	486	
FT	STRAND	487	497	
FT	TURN	503	505	
FT	STRAND	508	510	
FT	TURN	512	513	
FT	TURN	516	517	
FT	TURN	519	520	
FT	TURN	522	523	
FT	TURN	528	529	
FT	TURN	533	534	
FT	TURN	538	539	
FT	HELIX	541	542	
FT	HELIX	544	553	
FT	TURN	554	555	
FT	STRAND	559	563	
FT	HELIX	564	569	
FT	TURN	570	570	
FT	HELIX	578	581	
FT	TURN	582	582	
FT	TURN	585	586	
FT	STRAND	588	591	
FT	TURN	593	594	
FT	STRAND	597	599	
FT	TURN	600	601	
FT	TURN	603	605	
FT	STRAND	609	611	
FT	STRAND	615	619	
FT	HELIX	620	637	
FT	TURN	639	640	
FT	TURN	642	647	
FT	TURN	650	651	
FT	TURN	654	655	

Query Match 93.3%; Score 56; DB 1; Length 708;  
 Best Local Similarity 91.7%; Pred. No. 0.012;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRAVYVMAVG 12  
 |||||  
 DB 359 ARRAVYVCAVG 370

RESULT 2

TRFL\_HUMAN STANDARD; PRT; 711 AA.  
 ID TRFL\_HUMAN  
 AC P02788; O00756; Q16780; Q16785; Q16786; Q16789; Q96KZ4; Q96KZ5;  
 AC Q9H123;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Lactoferrin precursor (lactoferrin) [Contains: Lactoferrin A,  
 Lactoferrin B; Lactoferrin C].  
 GN LTF OR LTF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=90384839; PubMed=2402455;  
 RA Rey M.W., Moloshuk S.L., de Boer H.A., Pieper F.R.;  
 RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
 RL Nucleic Acids Res. 18:5288-5288(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Cho Y.Y.;  
 RL Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Connolly C.M.;  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Liang Q., Jimenez-Flores R., Richardson T.;  
 RT "Molecular cloning and sequence analysis of human lactoferrin.";  
 RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RA Wei X., Han J., Rado T.A.;  
 RT "Human neutrophil lactoferrin coding and 5' flanking region DNA  
 sequences.";  
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Cheng H., Chen X., Huan L.;  
 RT "cDNA cloning and sequence analysis of human lactoferrin.";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Prostate;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueffing T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Millhys S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Pahay J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson W.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Matra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [8]  
 RP SEQUENCE OF 3-711 FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=90326549; PubMed=2374734;  
 RT Powell M.J., Ogden J.E.;  
 RL "Nucleotide sequence of human lactoferrin cDNA.";  
 RN [9]  
 RP SEQUENCE OF 20-711.  
 RX MEDLINE=85076667; PubMed=6510420;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
 RA Legend D., Spik G., Montreuil J., Jolles P.;  
 RT "Human lactoferrin: amino acid sequence and structural  
 comparisons with other transferrins.";  
 RL Eur. J. Biochem. 145:659-666(1984).  
 RN [10]  
 RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
 RX MEDLINE=82046817; PubMed=6794640;  
 RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "The present state of the human lactoferrin sequence. Study and  
 alignment of the cyanogen bromide fragments and characterization of  
 N- and C-terminal domains.";  
 RL Biochim. Biophys. Acta 670:243-254(1981).  
 RN [11]  
 RP SEQUENCE OF 609-711.  
 RX MEDLINE=82262043; PubMed=7049727;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "An 88 amino acid long C-terminal sequence of human  
 lactoferrin.";  
 RL FEBS Lett. 142:107-110(1982).  
 RN [12]  
 RP SEQUENCE OF 436-711 FROM N.A.  
 RX MEDLINE=88001031; PubMed=3477300;  
 RA Rado T.A., Wei X., Benz E.J. Jr.;  
 RT "Isolation of lactoferrin cDNA from a human myeloid library and  
 expression of mRNA during normal and leukemic myelopoiesis.";  
 RL Blood 70:989-993(1987).  
 RN [13]  
 RP SEQUENCE OF 237-711 FROM N.A.  
 RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
 RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schut K.,  
 RA Gao L., La Bastide M., Kaplan N., Greco L.H., Tang M., Porcel B.M.,  
 RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
 RA Dragan V., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
 RA Diaz-Perez S., Zhou X., Yu Y., Matanabe M., Doggett N., Garcia D.,  
 RA Sagripanti J.L.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
 RX MEDLINE=90064528; PubMed=2585506;  
 RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
 RT "Structure of human lactoferrin: crystallographic structure analysis  
 and refinement at 2.8-A resolution.";  
 RL J. Mol. Biol. 209:711-734(1989).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RA Hardas M., Anderson B.F., Baker E.N.;  
 RT "Structure of human dimeric lactoferrin refined at 2.2-A  
 resolution.";  
 RL Acta Crystallogr. D 51:629-646(1995).

[16] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
 RP MEDLINE=97156796; PubMed=9003186;  
 RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
 Baker E.N.;  
 RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
 binding properties and crystal structure of the histidine-  
 RT 253-->methionine mutant.";  
 RL Biochemistry 36:341-346(1997).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=99190892; PubMed=10089347;  
 RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
 RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
 RT saamori.";  
 RL Acta Crystallogr. D 55:403-407(1999).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=99192677; PubMed=10089508;  
 RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
 RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
 RT and analysis of ligand-induced conformational change.";  
 RL Acta Crystallogr. D 54:1319-1335(1998).  
 RN [19]  
 RP CHARACTERIZATION OF LACTOFERROXINS.  
 RX MEDLINE=91166929; PubMed=1369293;  
 RA Tani F., Ito K., Chiba H., Yoshikawa M.;  
 RT "Isolation and characterization of opioic antagonist peptides derived  
 RT from human lactoferrin.";  
 RL Agric. Biol. Chem. 54:1803-1810(1990).  
 RN [20]  
 RP VARIANTS THR-30 AND ARG-48.  
 RX MEDLINE=99091914; PubMed=9873069;  
 RA Kintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
 Cumshey M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
 Sugai J., Kumamachiavel G., Munier F., Schneider D.F.,  
 El Metri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
 Hejmanek J.F., Teng C.T.;  
 RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
 RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
 RL Mol. Vision 4:11-32(1998).  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- FUNCTION: LACTOFERROXIN A, B AND C HAVE OPIOID ANTAGONIST  
 CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
 CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
 CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: Composed of two homologous domains.  
 CC -1- SIMILARITY: Belongs to the transferrin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X51961; CAA37914.1; -  
 DR EMBL: U07643; AAB60324.1; -  
 DR EMBL: M93150; AAB6159.1; -  
 DR EMBL: M83202; AAA59511.1; -  
 DR EMBL: M83205; AAA58656.1; -

Query Match 93.3%; Score 56; DB 1; Length 711;

Best Local Similarity 91.7%; Pred No. 0.012; Mismatches 1; Indels 0; Gaps 0;

Matches 11; Conservative 0; Yismatches 1; Indels 0; Gaps 0;

1 ARRARVWAVG 12

Db 360 ARRARVWCAVG 371  
 RESULT 3  
 TRFL HORSE STANDARD; PRT; 695 AA.  
 ID TRFL HORSE  
 AC 07811;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-OCT-2003 (Rel. 42, Last annotation update)  
 DE Lactotransferrin precursor (lactoferrin) (Fragment).  
 GN LTF.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RT "cDNA sequence of mare lactoferrin.";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RC TISSUE=Milk;  
 RX MEDLINE=99296631; PubMed=10366507;  
 RA Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RT "Three-dimensional structure of mare ferric lactoferrin at 2.6-A  
 RT resolution.";  
 RL J. Mol. Biol. 289:303-317(1999).  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: Composed of two homologous domains.  
 CC -1- SIMILARITY: Belongs to the transferrin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ010930; CAA09407.1; -  
 DR PDB: 1B1X; 02-DEC-98.  
 DR PDB: 1B7U; 02-FEB-99.  
 DR PDB: 1B7Z; 02-FEB-99.  
 DR PDB: 1F9E; 10-FEB-01.  
 DR PDB: 1I6B; 13-FEB-02.  
 DR PDB: 1QW4; 14-JAN-00.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; Transferrin; 2.  
 DR PRINTS: PR00422; TRANSFERRIN.  
 DR SMART: SM00094; TR\_FER; 2.  
 DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE: PS00206; TRANSFERRIN\_2; 1.  
 DR PROSITE: PS00207; TRANSFERRIN\_3; 1.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal; 3D-structure.  
 FT NOW TER 1 1  
 FT SIGNAL <1 6  
 FT CHAIN 7 695 LACTOTRANSFERRIN.  
 FT REPEAT 7 350 1.  
 FT REPEAT 351 695 2.  
 FT DISULFID 15 51  
 FT DISULFID 25 42  
 FT DISULFID 121 204  
 FT DISULFID 163 179  
 FT DISULFID 166 189  
 FT DISULFID 176 187  
 FT DISULFID 237 251



FT DISULFID 354 386  
FT DISULFID 364 377  
FT DISULFID 411 690  
FT DISULFID 431 653  
FT DISULFID 463 538  
FT DISULFID 487 681  
FT DISULFID 497 511  
FT DISULFID 508 521  
FT DISULFID 579 593  
FT DISULFID 631 636  
FT METAL 66 66  
FT METAL 98 98  
FT METAL 198 198  
FT METAL 259 259  
FT METAL 401 401  
FT METAL 439 439  
FT METAL 532 532  
FT METAL 601 601  
FT BINDING 123 123  
FT BINDING 127 127  
FT BINDING 129 129  
FT BINDING 130 130  
FT BINDING 465 465  
FT BINDING 469 469  
FT BINDING 472 472  
FT CARBOHYD 472 472  
FT CARBOHYD 143 143  
FT CARBOHYD 287 287  
FT CARBOHYD 482 482  
FT STRAND 12 16  
FT HELIX 19 34  
FT TURN 35 36  
FT STRAND 40 44  
FT HELIX 48 56  
FT TURN 57 58  
FT STRAND 62 65  
FT HELIX 67 74  
FT TURN 76 78  
FT STRAND 80 88  
FT STRAND 95 95  
FT STRAND 97 105  
FT TURN 106 107  
FT HELIX 112 114  
FT TURN 116 117  
FT STRAND 119 122  
FT TURN 125 126  
FT TURN 128 131  
FT HELIX 132 137  
FT TURN 139 142  
FT HELIX 147 148  
FT STRAND 151 158  
FT HELIX 151 163  
FT STRAND 165 166  
FT TURN 169 171  
FT HELIX 173 175  
FT TURN 176 178  
FT TURN 183 187  
FT TURN 191 192  
FT TURN 194 195  
FT HELIX 197 206  
FT TURN 207 208  
FT STRAND 212 216  
FT HELIX 219 223  
FT HELIX 227 230  
FT TURN 231 232  
FT STRAND 233 237  
FT TURN 238 240  
FT STRAND 241 244  
FT HELIX 245 247  
FT TURN 248 250  
FT STRAND 254 257  
FT STRAND 260 264  
FT HELIX 270 284

IRON 1.  
IRON 1.  
IRON 1.  
IRON 1.  
IRON 2.  
IRON 2.  
IRON 2.  
IRON 2.  
CARBONATE 1.  
CARBONATE 1.  
CARBONATE 1 (VIA AMIDE NITROGEN).  
CARBONATE 1 (VIA AMIDE NITROGEN).  
CARBONATE 2.  
CARBONATE 2.  
CARBONATE 2 (VIA AMIDE NITROGEN).  
CARBONATE 2 (VIA AMIDE NITROGEN).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).

FT TURN 286 287  
FT TURN 294 295  
FT TURN 308 309  
FT STRAND 312 315  
FT HELIX 322 326  
FT HELIX 328 336  
FT TURN 337 338  
FT HELIX 341 349  
FT STRAND 351 356  
FT HELIX 358 370  
FT TURN 371 372  
FT STRAND 375 380  
FT HELIX 383 392  
FT TURN 393 393  
FT STRAND 397 400  
FT HELIX 402 410  
FT TURN 411 412  
FT STRAND 414 421  
FT HELIX 429 433  
FT STRAND 439 446  
FT HELIX 454 456  
FT TURN 458 459  
FT STRAND 462 464  
FT TURN 467 468  
FT TURN 470 473  
FT HELIX 474 484  
FT HELIX 489 491  
FT TURN 492 492  
FT TURN 496 497  
FT STRAND 499 500  
FT TURN 503 504  
FT TURN 506 507  
FT STRAND 508 508  
FT TURN 515 516  
FT TURN 520 521  
FT TURN 525 526  
FT TURN 528 529  
FT HELIX 531 540  
FT TURN 541 542  
FT STRAND 546 550  
FT HELIX 551 555  
FT TURN 556 557  
FT TURN 559 560  
FT HELIX 565 568  
FT TURN 569 569  
FT STRAND 572 573  
FT TURN 575 578  
FT STRAND 580 581  
FT TURN 584 586  
FT HELIX 587 592  
FT STRAND 596 598  
FT STRAND 602 605  
FT TURN 607 609  
FT HELIX 610 624  
FT TURN 626 627  
FT TURN 629 630  
FT HELIX 631 634  
FT TURN 637 638  
FT TURN 641 642

Query Match 85.0%; Score 51; DB 1; Length 695;  
Best Local Similarity 83.3%; Pred. No. 0.094; 2; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARRAVYNAVG 12  
DB 346 ARRRVYNAVG 357

RESULT 4  
TRFL\_PIG STANDARD; PRT; 704 AA.  
ID TRFL\_PIG  
AC P14652; Q23557;

DT 01-APR-1990 (Rel. 14, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DB Lactoferrin precursor (lactoferrin).  
 LN LFP.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 NCBI\_TaxID=9823;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92379101; PubMed=1511015;  
 RA Lyndon J.P., O'Malley B.R., Saucedo O., Lee T., Heaton D.R.,  
 RT "Nucleotide and primary amino acid sequence of porcine lactoferrin.";  
 RL Biochem. Biophys. Acta 1132:97-99(1992).  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92367939; PubMed=1503259;  
 RA Alexander L.J., Levine W.B., Teng C.T., Beattie C.W.;  
 RT "Cloning and sequencing of the porcine lactoferrin cDNA.";  
 RL Ann. Genet. 23:251-256(1992).  
 RN  
 RP SEQUENCE OF 20-49.  
 RX MEDLINE=90105538; PubMed=2605266;  
 RA Hutchens T.M., Magnusson J.S., Yip T.-T.;  
 RT "Rapid purification of porcine colostrum whey lactoferrin by affinity  
 chromatography on single-stranded DNA-agarose. Characterization,  
 RT amino acid composition and N-terminal amino acid sequence.";  
 RL Biochem. Biophys. Acta 999:323-329(1989).  
 CC  
 CC FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC  
 CC SUBUNIT: Monomer.  
 CC  
 CC SUBCELLULAR LOCATION: Secreted.  
 CC  
 CC DOMAIN: Composed of two homologous domains.  
 CC  
 CC SIMILARITY: Belongs to the transferrin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: M92089; AAA31102.1; -;  
 DR EMBL: M91327; AAA31059.1; -;  
 DR PIR: A45543; A45543.  
 DR HSP: 077698; ICE2.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; transferrin; 2.  
 DR PRINTS: PRO0422; TRANSFERRIN.  
 DR SMART: SM00094; TR\_FER; 2.  
 DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE: PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE: PS00207; TRANSFERRIN\_3; 1.  
 KM Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 704  
 FT REPEAT 20 359  
 FT REPEAT 360 704  
 FT DISULFID 28 62  
 FT DISULFID 38 53  
 FT DISULFID 130 213  
 FT DISULFID 172 188  
 FT DISULFID 185 196  
 FT DISULFID 246 260  
 FT DISULFID 363 395  
 FT DISULFID 373 386  
 FT DISULFID 420 699  
 FT DISULFID 472 547  
 BY SIMILARITY.  
 LACTOTRANSFERRIN.

FT DISULFID 496 690 BY SIMILARITY.  
 FT DISULFID 506 520 BY SIMILARITY.  
 FT DISULFID 517 530 BY SIMILARITY.  
 FT DISULFID 588 602 BY SIMILARITY.  
 FT DISULFID 640 645 BY SIMILARITY.  
 FT METAL 77 77 IRON 1 (BY SIMILARITY).  
 FT METAL 107 107 IRON 1 (BY SIMILARITY).  
 FT METAL 207 207 IRON 1 (BY SIMILARITY).  
 FT METAL 268 268 IRON 1 (BY SIMILARITY).  
 FT METAL 410 410 IRON 2 (BY SIMILARITY).  
 FT METAL 448 448 IRON 2 (BY SIMILARITY).  
 FT METAL 541 541 IRON 2 (BY SIMILARITY).  
 FT METAL 610 610 IRON 2 (BY SIMILARITY).  
 FT BINDING 132 132 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 136 136 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 138 138 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 139 139 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 474 474 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 478 478 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 480 480 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 481 481 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 491 491 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 12 12 G -> W (IN REF. 1).  
 FT CONFLICT 46 48 RRT -> TTR (IN REF. 3).  
 FT CONFLICT 51 51 M -> I (IN REF. 1).  
 FT CONFLICT 85 85 D -> G (IN REF. 1).  
 FT CONFLICT 121 121 MISSING (IN REF. 2).  
 FT CONFLICT 132 132 T -> I (IN REF. 1).  
 FT CONFLICT 132 132 E -> S (IN REF. 1).  
 FT CONFLICT 284 284 E -> Q (IN REF. 1).  
 FT CONFLICT 573 573 D -> N (IN REF. 1).  
 FT CONFLICT 590 590 D -> N (IN REF. 1).  
 FT CONFLICT 625 625 V -> M (IN REF. 1).  
 FT CONFLICT 662 662 V -> C (IN REF. 1).  
 FT CONFLICT 686 704 NLKOSVSPFLGCAFMFR -> T (IN REF. 1).  
 SQ SEQUENCE 704 AA; 77625 MW; 93261EFP608AD358 CRC64;  
 Query Match 81.7%; Score 49; DB 1; Length 704;  
 Best Local Similarity 75.0%; Pred. No. 0.22;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 ARRAYVMAVG 12  
 Db 355 ARQKVTWCAVG 366  
 RESULT 5  
 TFFL CAMDR STANDARD; PRT; 708 AA.  
 ID TFFL CAMDR  
 AC Q9TUM0; O9WZS5;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Lactotransferrin precursor (lactoferrin).  
 GN LFP.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 NCBI\_TaxID=9638;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R., Ackermann M., Farah Z., Puhon Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";  
 RL Int. Dairy J. 9:481-486(1999).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -1- SUBUNIT: Monomer (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: Composed of two homologous domains.  
 CC -1- SIMILARITY: Belongs to the transferrin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ131674; CAB53387.1; -.  
 DR EMBL: AF165879; AAP82241.1; -.  
 DR PDB: 1DTZ; 2O-JUN-01.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; Transferrin; 2.  
 DR PRINTS: PRO0422; TRANSFERRIN.  
 DR SMART: SM00094; TR\_FER; 2.  
 DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE: PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE: PS00207; TRANSFERRIN\_3; 2.  
 DR PROSITE: PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal; 3D-structure.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.  
 FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 ION 1 (BY SIMILARITY).  
 FT METAL 111 111 ION 1 (BY SIMILARITY).  
 FT METAL 211 211 ION 1 (BY SIMILARITY).  
 FT METAL 272 272 ION 1 (BY SIMILARITY).  
 FT METAL 414 414 ION 2 (BY SIMILARITY).  
 FT METAL 452 452 ION 2 (BY SIMILARITY).  
 FT METAL 545 545 ION 2 (BY SIMILARITY).  
 FT METAL 614 614 ION 2 (BY SIMILARITY).  
 FT BINDING 136 136 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 140 140 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 142 142 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 FT SIMILARITY).  
 FT BINDING 143 143 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 FT SIMILARITY).  
 FT BINDING 478 478 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 482 482 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 484 484 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT SIMILARITY).  
 FT BINDING 485 485 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 261 261 F -> S (IN REF. 2).  
 FT CONFLICT 304 304 G -> A (IN REF. 2).  
 FT CONFLICT 330 330 S -> P (IN REF. 2).  
 FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).  
 FT CONFLICT 506 506 L -> F (IN REF. 2).  
 FT CONFLICT 609 609 A -> P (IN REF. 2).  
 FT CONFLICT 642 642 R -> Q (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;  
 Query Match 80.0%; Score 48; DB 1; Length 708;  
 Best Local Similarity 81.8%; Pred. No. 0.34;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 RRARYMAVG 12  
 DB 360 RRAQYVCAVG 370  
 RESULT 6  
 ID TRFL\_BOVIN STANDARD; PRT; 708 AA.  
 AC TRFL\_BOVIN Q29629; Q9MZV3;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin) [contains: Lactoferricin B  
 DE (LfCIN B)].  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Submaxillary gland;  
 RX MEDLINE=91160550; PubMed=2001696;  
 RA Pierce A., Colavizza D., Benalissa M., Maes P., Tartar A.,  
 RA Montreuil J., Spik G.;  
 RT "Molecular cloning and sequence analysis of bovine lactotransferrin.";  
 RL Eur. J. Biochem. 196;1177-184(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92028986; PubMed=1718281;  
 RA Goodman R.E., Schanbacher F.L.;  
 RA "Bovine lactoferrin mRNA: sequence, analysis, and expression in the  
 RT mammary gland.";  
 RL Biochem. Biophys. Res. Commun. 180;75-84(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Teag T.C., Burns D.K., Wang F., Pan Y.C.E., Schmidt A.M., Stern D.;  
 RT "Cloning of a 80-kD advanced glycosylation end product (AGE) binding  
 RL protein from bovine lung.";  
 RN FASEB J. 6;233-233(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood, and Mammary gland;  
 RX MEDLINE=94466164; PubMed=8206385;  
 RA Seyfert H.-M., Tuckovitz A., Interthal H., Kozzan D., Hobom G.;  
 RT "Structure of the bovine lactoferrin-encoding gene and its promoter.";  
 RL Gene 143;265-269(1994).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Nakamura I., Shimazaki K., Yagi Y., Watanabe A.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE OF 20-59.  
 RX MEDLINE=90031466; PubMed=2805645;  
 RA Rejman U.J., Hegarty H.W., Hurley W.L.;  
 RT "Purification and characterization of bovine lactoferrin from  
 RT secretions of the involuting mammary gland: identification of  
 RT multiple molecular weight forms.";

Comp. Biochem. Physiol. 93B:929-934 (1989).  
[7]  
RX X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RX MEDLINE=98062367; PubMed=9398529;  
RA Moore S.A., Anderson B.F., Groen C.R., Haridas M., Baker E.N.;  
RT "Three-dimensional structure of ferric bovine lactoferrin at 2.8-A  
resolution";  
RL J. Mol. Biol. 274:222-236(1997).  
[8]  
RX STRUCTURE BY NMR OF 36-60.  
RX MEDLINE=98190007; PubMed=9521752;  
RA Huang P.M., Zhou N., Shan X., Atrowsmith C.H., Vogel H.J.;  
RT "Three-dimensional solution structure of lactoferricin B, an  
antimicrobial peptide derived from bovine lactoferrin";  
RL Biochemistry 37:4288-4298(1998).  
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
OF AN ANION, USUALLY BICARBONATE.  
CC -1- FUNCTION: LACTOFERRICIN B IS AN ANTIMICROBIAL PEPTIDE.  
CC -1- SUBUNIT: Monomer.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- DOMAIN: Composed of two homologous domains.  
CC -1- SIMILARITY: Belongs to the transferrin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL: X57084; CAA0366.1; -;  
DR EMBL: M63502; AAA30617.1; -;  
DR EMBL: L08604; AAA30609.1; -;  
DR EMBL: L19993; AAA21722.1; -;  
DR EMBL: L19982; AAA21722.1; JOINED.  
DR EMBL: L19983; AAA21722.1; JOINED.  
DR EMBL: L19984; AAA21722.1; JOINED.  
DR EMBL: L19985; AAA21722.1; JOINED.  
DR EMBL: L19986; AAA21722.1; JOINED.  
DR EMBL: L19988; AAA21722.1; JOINED.  
DR EMBL: L19989; AAA21722.1; JOINED.  
DR EMBL: L19990; AAA21722.1; JOINED.  
DR EMBL: L19991; AAA21722.1; JOINED.  
DR EMBL: L19992; AAA21722.1; JOINED.  
DR EMBL: A8046664; BAB03470.1; -;  
DR PIR: I4519; TFEOL.  
DR PDB: 1LFC; 03-DEC-97.  
DR GlycoSiteDB: P24627; -;  
DR InterPro: IPR001156; Transferrin.  
DR Pfam: PF00405; transferrin; 2.  
DR PRINTS: PRO0422; TRANSFERRIN.  
DR SMART: SM00094; TR\_FER; 2.  
DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE: PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE: PS00207; TRANSFERRIN\_3; 2.  
KM Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
KW Signal; Antibiotic; 3D-structure.  
FT CHAIN 1 708 LACTOTRANSFERRIN.  
FT PEPTIDE 36 60 LACTOFERRICIN B.  
FT REPEAT 20 363 1.  
FT DISULFID 364 708 2.  
FT DISULFID 28 64 BY SIMILARITY.  
FT DISULFID 38 55  
FT DISULFID 134 217 BY SIMILARITY.  
FT DISULFID 176 192 BY SIMILARITY.  
FT DISULFID 189 200 BY SIMILARITY.  
FT DISULFID 250 264 BY SIMILARITY.  
FT DISULFID 367 399 BY SIMILARITY.  
FT DISULFID 377 390 BY SIMILARITY.

FT DISULFID 424 703 BY SIMILARITY.  
FT DISULFID 444 666 BY SIMILARITY.  
FT DISULFID 476 551 BY SIMILARITY.  
FT DISULFID 500 694 BY SIMILARITY.  
FT DISULFID 510 524 BY SIMILARITY.  
FT DISULFID 521 534 BY SIMILARITY.  
FT DISULFID 592 606 BY SIMILARITY.  
FT DISULFID 644 649 BY SIMILARITY.  
FT METAL 79  
FT METAL 111 111 IRON 1.  
FT METAL 211 211 IRON 1.  
FT METAL 272 272 IRON 1.  
FT METAL 414 414 IRON 2.  
FT METAL 452 452 IRON 2.  
FT METAL 545 545 IRON 2.  
FT METAL 614 614 IRON 2.  
FT BINDING 136 136 CARBONATE 1.  
FT BINDING 140 140 CARBONATE 1.  
FT BINDING 142 142 CARBONATE 1.  
FT BINDING 143 143 CARBONATE 1.  
FT BINDING 148 148 CARBONATE 2.  
FT BINDING 478 478 CARBONATE 2.  
FT BINDING 482 482 CARBONATE 2 (VIA AMIDE NITROGEN).  
FT BINDING 484 484 CARBONATE 2 (VIA AMIDE NITROGEN).  
FT BINDING 485 485 CARBONATE 2 (VIA AMIDE NITROGEN).  
FT CARBOHYD 252 252  
FT CARBOHYD 387 387  
FT CARBOHYD 495 495  
FT CARBOHYD 564 564  
FT CARBOHYD 564 564  
FT CONFLICT 63 63  
FT CONFLICT 66 67  
FT CONFLICT 145 145  
FT CONFLICT 164 165  
FT CONFLICT 264 264  
FT CONFLICT 273 273  
FT CONFLICT 281 281  
FT CONFLICT 291 291  
FT CONFLICT 297 297  
FT CONFLICT 340 340  
FT CONFLICT 418 418  
FT CONFLICT 439 439  
FT CONFLICT 459 459  
FT CONFLICT 514 514  
FT CONFLICT 632 632  
FT STRAND 25 29  
FT HELIX 32 45  
FT HELIX 46 48  
FT STRAND 53 57  
FT HELIX 61 69  
FT TURN 70 71  
FT STRAND 75 78  
FT HELIX 80 87  
FT TURN 89 91  
FT STRAND 93 102  
FT STRAND 107 108  
FT STRAND 110 118  
FT TURN 126 127  
FT TURN 129 130  
FT STRAND 133 135  
FT TURN 138 139  
FT TURN 141 144  
FT HELIX 145 155  
FT TURN 156 156  
FT TURN 159 161  
FT HELIX 164 169  
FT TURN 170 171  
FT STRAND 175 176  
FT TURN 178 179  
FT TURN 182 184  
FT HELIX 186 188  
FT TURN 190 191  
FT TURN 191 191  
RA -> A (IN REF. 4).  
I -> V (IN REF. 1 AND 4).  
LQ -> PP (IN REF. 1).  
C -> Y (IN REF. 4).  
A -> P (IN REF. 4).  
G -> A (IN REF. 4).  
S -> R (IN REF. 4).  
F -> S (IN REF. 2).  
G -> A (IN REF. 1).  
I -> V (IN REF. 4).  
H -> Y (IN REF. 1).  
K -> R (IN REF. 4).  
A -> R (IN REF. 1).  
H -> R (IN REF. 5).

FT TURN 196 197

Query Match 75.0%; Score 45; DB 1; Length 708;  
 Best Local Similarity 75.0%; Pred. No. 1.2;  
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
 |||||  
 DB 359 ARTRVWCAVG 370

RESULT 7  
 TRFL\_MOUSE STANDARD; PRT; 708 AA.

AC Q29477, Q29479, 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OC NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee Y., Yu S., Kim S., Lee K., Yu D.;  
 RA Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=94380047; Pubmed=8093048;  
 RA Le Provost F., Nocart M., Guerin G., Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 RT relevant locus to bovine U2 synteny group."  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: Composed of two homologous domains.  
 CC -1- SIMILARITY: Belongs to the transferrin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U53857; AAA97958.1; -;  
 DR EMBL: X78902; CA55517.1; -;  
 DR HSSP: O77696; ICE2.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; transferrin; 2.  
 DR PRINTS: PRO0422; TRANSFERRIN.  
 DR SMART: SM00094; TR\_FER; 2.  
 DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE: PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE: PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 RT

FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.  
 FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 BY SIMILARITY.  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 142 142 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 143 143 SIMILARITY).  
 FT BINDING 143 143 CARBONATE 1 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 478 478 SIMILARITY).  
 FT BINDING 482 482 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 484 484 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT BINDING 485 485 SIMILARITY).  
 FT BINDING 485 485 CARBONATE 2 (VIA AMIDE NITROGEN) (BY  
 FT CARBOHYD 252 252 SIMILARITY).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 56 56 I -> V (IN REF. 2).  
 FT CONFLICT 88 88 L -> R (IN REF. 2).  
 FT CONFLICT 124 124 Q -> K (IN REF. 2).  
 FT CONFLICT 154 154 F -> P (IN REF. 2).  
 FT CONFLICT 304 304 S -> R (IN REF. 2).  
 FT CONFLICT 414 414 D -> G (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77358 MW; FEDA5C83539960D CEC64;

Query Match 73.3%; Score 44; DB 1; Length 708;  
 Best Local Similarity 75.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12  
 |||||  
 DB 359 ARTRVWCAVG 370

RESULT 8  
 TRFL\_MOUSE STANDARD; PRT; 707 AA.

AC P08071, P70690, Q61799, Q922P2;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RX MEDLINE=87280033; Pubmed=3611056;  
 RA Pentecost B.T., Teng C.T.;  
 RT "Lactotransferrin is the major estrogen inducible protein of mouse

RT uterine secretions.";  
 RL J. Biol. Chem. 262:10134-10139(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RA Morishita K.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22386257; PubMed=12477932;  
 RA Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.T.,  
 RA Rana S.S., Loughran N.A., Peters G.O., Abramson R.D., Mullen S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatane P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RX MEDLINE=92042099; PubMed=1939212;  
 RA Liu Y., Teng C.T.;  
 RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";  
 RL J. Biol. Chem. 266:21880-21885(1991).  
 CC -1- FUNCTION: TRANSPORTERS ARE IRON BINDING PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: Composed of two homologous domains.  
 CC -1- SIMILARITY: Belongs to the transferrin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: J03298; AAA40525.1; -;  
 DR EMBL: DB8510; BAA13633.1; -;  
 DR EMBL: BC006904; AA006904.1; -;  
 DR EMBL: M74778; AAA39427.1; -;  
 DR HSP: P02788; ICB6.  
 DR MGD: MGI:96837; Lcf.  
 DR InterPro: IPR001156; Transferrin.  
 DR Pfam: PF00405; Transferrin; 2.  
 DR PRINTS: PR00422; TRANSFERRIN.  
 DR SMART: SMO0094; TR\_FER; 2.  
 DR PROSITE: PS00205; TRANSFERRIN 1; 1.  
 DR PROSITE: PS00206; TRANSFERRIN 2; 2.  
 DR PROSITE: PS00207; TRANSFERRIN 3; 2.  
 KM Transpore; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KM Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 707 LACTOTRANSFERRIN.  
 FT REPEAT 20 357 1.  
 FT REPEAT 358 707 2.  
 FT DISULFID 27 63 BY SIMILARITY.  
 FT DISULFID 37 54 BY SIMILARITY.

FT DISULFID 133 216 BY SIMILARITY.  
 FT DISULFID 175 191 BY SIMILARITY.  
 FT DISULFID 188 199 BY SIMILARITY.  
 FT DISULFID 249 263 BY SIMILARITY.  
 FT DISULFID 366 398 BY SIMILARITY.  
 FT DISULFID 376 389 BY SIMILARITY.  
 FT DISULFID 423 702 BY SIMILARITY.  
 FT DISULFID 443 665 BY SIMILARITY.  
 FT DISULFID 475 550 BY SIMILARITY.  
 FT DISULFID 499 693 BY SIMILARITY.  
 FT DISULFID 509 523 BY SIMILARITY.  
 FT DISULFID 520 533 BY SIMILARITY.  
 FT DISULFID 591 605 BY SIMILARITY.  
 FT DISULFID 643 648 BY SIMILARITY.  
 FT METAL 78 78 IRON 1 (BY SIMILARITY).  
 FT METAL 110 110 IRON 1 (BY SIMILARITY).  
 FT METAL 210 210 IRON 1 (BY SIMILARITY).  
 FT METAL 271 271 IRON 1 (BY SIMILARITY).  
 FT METAL 413 413 IRON 2 (BY SIMILARITY).  
 FT METAL 451 451 IRON 2 (BY SIMILARITY).  
 FT METAL 544 544 IRON 2 (BY SIMILARITY).  
 FT METAL 613 613 IRON 2 (BY SIMILARITY).  
 FT BINDING 135 135 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 139 139 CARBONATE 1 (BY SIMILARITY).  
 FT BINDING 141 141 CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY).  
 FT BINDING 142 142 CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY).  
 FT BINDING 477 477 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 481 481 CARBONATE 2 (BY SIMILARITY).  
 FT BINDING 483 483 CARBONATE 2 (VIA AMIDE NITROGEN) (BY SIMILARITY).  
 FT BINDING 484 484 CARBONATE 2 (VIA AMIDE NITROGEN) (BY SIMILARITY).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1 2 MR -> IOG (IN REF. 1).  
 FT CONFLICT 25 25 R -> Q (IN REF. 2).  
 FT CONFLICT 82 82 M -> L (IN REF. 2).  
 FT CONFLICT 359 359 S -> T (IN REF. 2).  
 FT CONFLICT 382 382 A -> D (IN REF. 1).  
 FT CONFLICT 449 449 E -> G (IN REF. 2).  
 FT CONFLICT 629 629 L -> V (IN REF. 1).  
 SQ SEQUENCE 707 AA; 77865 MW; F26AE0340AC19A8 CRC64;  
 Query Match 71.7%; Score 43; DB 1; Length 707;  
 Best Local Similarity 66.7%; Pred. No. 2.8;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Cx 1 ARARVWMAVG 12  
 Db 358 ASKARVWCAVG 369  
 RESULT 9  
 CAHB\_HUMAN STANDARD; PRT; 328 AA.  
 ID CAHB\_HUMAN  
 AC 075493; O60596; O9UEC4;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Carbonic anhydrase-related protein 2 precursor (CARP-2) (CA-RP II)  
 DE (CA-XI) (Carbonic anhydrase-related protein 11) (CARP XI) (CA-RP XI).  
 GN CALL OR CARP2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CX NCBI\_TaxID=9606;  
 OK [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99097242; PubMed=9878543;  
 RA Bellingham J., Gregory-Evans K., Gregory-Evans C.Y.;  
 RT "Sequence and tissue expression of a novel human carbonic

RT anhydrase-related protein, CARP-2, mapping to chromosome 19q13.3.";  
 RL Biochem. Biophys. Res. Commun. 253:364-367(1998).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=99280326; PubMed=10350627;  
 RA Fujikawa-Adachi K., Nishimori I., Taguchi T., Yuri K., Onishi S.,  
 RT "cDNA sequence, mRNA expression, and chromosomal localization of human  
 carbonic anhydrase-related protein, CA-RP XI.";  
 RL Biochim. Biophys. Acta 1431:518-524(1999).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallamy S.J.,  
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [4]  
 RP SEQUENCE OF 48-314 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99097349; PubMed=9878252;  
 RA Lovejoy D.A., Hewett-Emmett D., Porter C.A., Cepoi D., Sheffield A.,  
 RA Vale W.W., Tashian R.E.,  
 RT "Evolutionarily conserved, 'acatalytic' carbonic anhydrase-related  
 protein XI contains a sequence motif present in the neurotrophin  
 savignine: the human CA-RP XI gene (CA11) is embedded between the  
 RT secretor gene cluster and the DBP gene at 19q13.3.";  
 RL Genomics 54:484-493(1998).  
 CC -1- FUNCTION: Does not have a catalytic activity.  
 CC -1- SUBCELLULAR LOCATION: Secreted (potential).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN THE BRAIN WITH  
 CC MODERATE EXPRESSION ALSO PRESENT IN SPINAL CORD AND THYROID.  
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase  
 CC family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL/ AF067662; AAC09689.1; -;  
 CC EMBL/ AB018195; EAA36840.1; -;  
 CC EMBL/ BC002662; AA02662.1; -;  
 CC EMBL/ AF050106; AAD08802.1; -;  
 CC PIR: J03075; J03075.  
 CC HSSP: P00918; 1BV3.  
 CC GeneW, HGNC:1370; CA11.  
 CC MIM: 604644; -;  
 CC GO: GO:0004089; P:carbonate dehydratase activity; TAS.  
 CC InterPro: IPR001148; Euk Coanhd.  
 CC Pfam: PF00194; carb anhydrase; 1.  
 CC ProDom: PD000865; Euk Coanhd; 1.  
 CC PROSITE: PS00162; Euk\_CO2\_AHYDRASE; FALSE\_NEG.

KW Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 328 CARBONIC ANHYDRASE-RELATED PROTEIN 2.  
 FT CONFLICT 23 24 AH -> GN (IN REF. 2).  
 FT CONFLICT 75 75 L -> V (IN REF. 1).  
 FT CONFLICT 280 280 I -> M (IN REF. 4).  
 SQ SEQUENCE 328 AA; 36238 MW; AAFCA08718B2E857 CRC64;  
 Query Match 66.7%; Score 40; DB 1; Length 328;  
 Best Local Similarity 66.7%; Pred. No. 4.8;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 APPARVMAAVG 12  
 DB 9 APPALVMAALG 20  
 RESULT 10  
 C13A\_MYCTU STANDARD; PRT; 449 AA.  
 ID C13A\_MYCTU  
 AC 008447;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative cytochrome P450 135A1 (EC 1.14.-.-).  
 GN CYP135A1 OR RV0327C OR MT0342 OR MTCY63.32C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean A., Moutle S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.,  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RX MEDLINE=22206994; PubMed=12218036;  
 RA Fleischman R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 RA Kolony A.F., Nelson W.C., Umayam L.A., Enomoeva M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill U., Mikula A.,  
 RA Bishtai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.,  
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RL Bacteriol. 184:5479-5490(2002).  
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL/ Z96800; CAA09576.1; -;  
 CC EMBL/ AE006940; AAK4565.1; -;  
 CC PIR: H70526; H70526.  
 CC TIGR: MT0342; -;  
 CC TubercuList: RV0327C; -;  
 CC InterPro: IPR001128; Cytochrome\_P450.

```

DR Pfam: PF00067, P450, 1.
DR PRINTS: PR00385, P450.
DR PROSITE: PS00086; CYTOCHROME P450, 1.
KW Hypothetical protein; Oxidoreductase; Monooxygenase; Heme;
KW Complete proteome.
FT MISC 383 383 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 449 AA; 50010 MW; 19253CF1855EDCF5 CRC64;

Query Match
Best Local Similarity 70.0%; Score 40; DB 1; Length 449;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RARVVAAYG 12
DB 285 QSQVVAAYG 294

RESULT 11
REP2_ECOLI STANDARD; PRT; 285 AA.
AC P03066; Q47411, Created
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Replication initiation protein.
GN REPA OR REPAI.
OS Escherichia coli.
OC Plasmid IncFII R100, Plasmid IncFII R1, and Plasmid IncFII NR1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=81074309; PubMed=7003300;
RA Rosen J., Ryder T., Inokuchi H., Ohtsubo H., Ohtsubo E.;
RT "Genes and sites involved in replication and incompatibility of an
RT R100 plasmid derivative based on nucleotide sequence analysis.";
RL Mol. Gen. Genet. 179:527-537(1980).
RN [2]
RP SEQUENCE OF 1-77 FROM N.A.
RC PLASMID=IncFII R100, and IncFII R1;
RX MEDLINE=81173118; PubMed=6163994;
RA Rosen J., Ryder T., Ohtsubo H., Ohtsubo E.;
RT "Role of RNA transcripts in replication incompatibility and copy
RT number control in antibiotic resistance plasmid derivatives.";
RL Nature 290:794-797(1981).
RN [3]
RP SEQUENCE OF 270-285 FROM N.A.
RC PLASMID=IncFII R1;
RX MEDLINE=88289416; PubMed=3041379;
RA Masai H., Arai K.;
RT "Repa protein- and oriR-dependent initiation of R1 plasmid
RT replication: identification of a rho-dependent transcription
RT terminator required for cis-action of repA protein.";
RL Nucleic Acids Res. 16:6493-6514(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC PLASMID=IncFII NR1;
RX MEDLINE=89011975; PubMed=3050127;
RA Dong X., Womble D.D., Rownd R.H.;
RT "In vivo studies on the cis-acting replication initiator protein of
RT IncFII plasmid NR1.";
RL J. Mol. Biol. 202:495-509(1988).
RN [5]
RP SEQUENCE FROM N.A.
RC PLASMID=IncFII NR1;
RX MEDLINE=85160860; PubMed=2580909;
RA Womble D.D., Sampathkumar P., Easton A.M., Luckow V.A., Rownd R.H.;
RT "Transcription of the replication control region of the IncFII
RT J. Mol. Biol. 181:395-410(1985).
RN [6]

```

```

RP SEQUENCE FROM N.A.
RC PLASMID=IncFII R100;
RX MEDLINE=86319522; PubMed=3019092;
RA Ohtsubo H., Ryder T.B., Maeda Y., Armstrong K., Ohtsubo E.;
RT "DNA replication of the resistance plasmid R100 and its control.";
RL Adv. Biophys. 21:115-133(1986).
CC -1- FUNCTION: This protein is essential for plasmid replication; it is
CC involved in copy control functions.
CC -1- SIMILARITY: BELONGS TO THE INCFII REPA FAMILY.
CC -----
CC EMBL: J01762; AAA92257.1; -.
CC EMBL: J01770; -. NOT ANNOTATED_CDS.
CC EMBL: X12587; CAA31100.1; -.
CC EMBL: X12776; CAA31263.1; -.
CC EMBL: X02302; CAA26168.1; -.
CC EMBL: X02302; CAA26169.1; ALT_INIT.
CC EMBL: M26840; AAA26067.1; -.
CC PIR: A03602; IDECRP.
CC PIR: I64780; I64780.
CC InterPro: IPR003446; RepIctn.
DR Pfam: PF02387, IncFII_repa, 1.
DR Plasmid, DNA replication; Plasmid copy control.
FT CONFLICT 55 55 R -> H (in Ref. 6).
SQ SEQUENCE 285 AA; 32755 MW; A21C9D59D24B26B CRC64;

Query Match
Best Local Similarity 87.5%; Score 39; DB 1; Length 285;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AARARVAV 8
DB 168 AARSRVAV 175

RESULT 12
YH92_AERPE STANDARD; PRT; 176 AA.
AC Q9YB03;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein APEI792.
GN APEI792.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankei A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takemizu M., Masuda S., Tanahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki U., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
CC -1- SIMILARITY: Belongs to the UPF0097 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```





RA	Laird G.X., Langford C.F., Leyerstra M.A., Lloyd C., Lloyd D.M.,
RA	Matlay I.D., Masfrehri-Mohammadi M., Matthews L.H., Mccann O.T.,
RA	Mcclay J., McLaren S., McMurtry A.A., Milne S.A., Mortimore B.J.,
RA	Odell C.N., Pavitt R., Pearce A.V., Pearson S., Phillimore B.U.C.T.,
RA	Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA	Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA	Soderlund C., Spraggon L., Steward C.A., Sulston U.E., Swann R.M.,
RA	Vandin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA	Williams L., Williams S.A., Williamson H., Wilner T.E., Wilting L.,
RA	Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shintzu N.,
RA	Mitoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh U.,
RA	Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,
RA	Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA	Dorman A., Fang F., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA	Lewis J., Lewis S., Lin S.-P., Loh P., Malat E., Nguyen T., Pan H.,
RA	Pheon S., Qi S., Qian Y., Ray L., Ren Q., Shauli S., Sloan D., Song L.,
RA	Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA	Zhan M., Zhang G., Chisose S., Murray J., Miller R., Minx P.,
RA	Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA	Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins U.,
RA	Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA	Scheet P., Walker C., Wamsley A., Wohlmann P., Popin K., Nelson J.,
RA	Kort J., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA	Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA	Budart M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA	Edelmann L., Kim U.J., Shizuya A., Simon M.I., Dunamski J.P.,
RA	Peybard K.P., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA	O'Brien K.P., Wilkinson P., Bodenreich A., Hartman K., Hu X.,
RA	Khan A.S., Lane L., Tiliahun Y., Wright H.,
RL	"The DNA sequence of human chromosome 22.";
CC	Nature 402:485-495(1999).
CC	- FUNCTION: Receptor for the hormone galanin.
CC	- SUBCELLULAR LOCATION: Integral membrane protein.
CC	- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sb.ch/announce/">http://www.isb-sb.ch/announce/</a> or send an email to <a href="mailto:license@isb-sb.ch">license@isb-sb.ch</a> ).
CC	-----
DR	EMBL, AF073799; AAC35944.1; -
DR	EMBL, AF067733; AAC18860.1; -
DR	EMBL, AF129514; AAD47348.2; -
DR	EMBL, AF129513; AAD47348.2; COINED.
DR	EMBL, AB065934; BAC06149.1; -
DR	EMBL, 297630; CAB24831.1; -
DR	GeneW: HGNC:4134; GALT3.
DR	MM, 603692; -
DR	GO: GO:0016021; C:integral to membrane; TAS.
DR	GO: GO:0005886; C:plasma membrane; TAS.
DR	GO: GO:0004966; F:galanin receptor activity; TAS.
DR	GO: GO:0007631; P:feeding behavior; TAS.
DR	GO: GO:0007611; P:learning and/or memory; TAS.
DR	GO: GO:0007194; P:negative regulation of adenylate cyclase ac. . ; TAS.
DR	GO: GO:0007218; P:neuropeptide signaling pathway; TAS.
DR	GO: GO:0007268; P:synaptic transmission; TAS.
DR	InterPro: IPR000276; GPCR_Rhodopsn.
DR	Pfam: PFD0001; Tcm 1; 1.
DR	PRINTS: PRO0237; GPCR_RHODOPSN.
DR	PROSITE: PS00337; G PROTEIN RECP FL 1; 1.
DR	PROSITE: PS00362; G PROTEIN RECP FL 2; 1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;
KW	Multigene family; Lipoprotein; Palmitate.
FT	DOMAIN 1 20
FT	TRANSMEM 21 41
FT	DOMAIN 42 57
FT	TRANSMEM 58 78
FT	DOMAIN 79 96
FT	TRANSMEM 97 118
FT	DOMAIN 119 138
FT	CITOPLASTIC (POTENTIAL).

FT	TRANSMEM	139	159	4 (POTENTIAL).	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	184	160	5 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	185	205	6 (POTENTIAL). <td>EXTRACELLULAR (POTENTIAL).</td>	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	206	236	7 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	237	257	8 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	258	259	9 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	260	280	10 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	281	368	11 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	6	6	12 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	DISULFID	95	172	13 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	LIPID	308	308	14 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	218	225	15 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	310	318	16 (POTENTIAL). <td>CYTOPLASMIC (POTENTIAL).</td>	CYTOPLASMIC (POTENTIAL).
SEQ	SEQUENCE	368 AA;	39573 MW;	88FA6B6642C3150 CRC64;	POLY-ARG.
Query March		61.7%;	Score 37;	DB 1;	Length 368;
Best Local Similarity		63.6%;	Pred. No. 19;		
Matches	7;	Conservative	1;	Mismatches	3;
				Indels	0;
				Gaps	0;
QY	2 RRAVMAAVG 12				
Db	206 RRLFLMAAVG 216				
RESULT 15					
EGFR_HUMAN	STANDARD;	PRT;	651 AA.		
ID	P08236; Q96CL9;				
AC	01-AUG-1988 (Rel. 08, Created)				
DT	01-AUG-1988 (Rel. 08, last sequence update)				
DT	10-OCT-2003 (Rel. 42, last annotation update)				
DT	Beta-glucuronidase precursor (EC 3.2.1.31) (Beta-G1).				
GN	GUSB.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
CC	NCBI_TaxID=9606;				
RN	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.				
RP	TISSUE=Colon;				
RC	MEDLINE=22388257; PubMed=3468507;				
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,				
RA	Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Heien F.,				
RA	Ditcheenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stepien M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M., Usdin T.B., Toshiyuki S., Carrinchi P., Mullaly S.J.,				
RA	Rana S.S., Loguelfino N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., McMan P.J., McKenna K.J., Malek J.A., Gunaratne P.H.,				
RA	Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.U., Hilyk S.W.,				
RA	Villalón D.K., Muzny D.M., Sodersten E.J., Lu X., Gibbs R.A.,				
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield J.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,				
RA	Scherer A., Schein J.B., Jones S.U.W., Matra M.A.,				
RT	"Generation and initial analysis of more than 15,000 full-length				
RT	human and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).				
RN	[3]				
RP	SEQUENCE OF 1-70 FROM N.A.				
RX	MEDLINE=92009900; PubMed=1916806;				

RA Shipley J.M., Miller R.D., Wu B.M., Grubb J.H., Christensen S.G.,  
 RA Kyle J.W., Sly W.S., "Analysis of the 5' flanking region of the human beta-glucuronidase  
 RT gene.",  
 RL Genomics 10:1009-1018(1991).  
 [4]  
 RP SEQUENCE OF 520-585 FROM N.A.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=85232043; PubMed=924735;  
 RA Guise K.S., Korneluk R.G., Waye U., Lamhorwah A.-W., Quan F.,  
 RA Palmer R., Ganschow R.B., Sly W.S., Gravel R.A.,  
 RT "Isolation and expression in *Escherichia coli* of a cDNA clone encoding  
 RL human beta-glucuronidase.",  
 RL Gene 34:105-110(1985).  
 [5]  
 RP SEQUENCE OF 23-32 AND 160-175.  
 RC TISSUE=Placenta;  
 RX MEDLINE=92162201; PubMed=1311180;  
 RA Tanaka U., Gasa S., Sakurada K., Miyazaki T., Kasai M., Makita A.,  
 RT "Characterization of the subunits and sugar moiety of human placental  
 RL and leukemic beta-glucuronidase.",  
 RL Biol. Chem. Hoppe-Seyler 373:57-62(1992).  
 [6]  
 RP CARBOHYDRATE-LINKAGE SITE ASN-272.  
 RX MEDLINE=22660472; PubMed=12754519;  
 RA Zhang H., Li X.-J., Martin D.B., Aebbersold R.,  
 RT "Identification and quantification of N-linked glycoproteins using  
 RT hydrazide chemistry, stable isotope labeling and mass spectrometry.",  
 RL Nat. Biotechnol. 21:660-666(2003).  
 [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RX MEDLINE=96185449; PubMed=8599764;  
 RA Jain S., Drendel W.B., Chen Z.W., Mathews F.S., Sly W.S., Grubb J.H.,  
 RT "Structure of human beta-glucuronidase reveals candidate lysosomal  
 RT targeting and active-site motifs.",  
 RL Nat. Struct. Biol. 3:375-381(1996).  
 [8]  
 RP VARIANT MPS-VII TRP-216.  
 RX MEDLINE=94154730; PubMed=8111412;  
 RA Verwoort R., Lissens W., Liebaers I.,  
 RT "Molecular analysis of a patient with hydrops fetalis caused by beta-  
 RT glucuronidase deficiency and evidence for additional pseudogenes.",  
 RL Hum. Mutat. 2:443-445(1993).  
 [9]  
 RP VARIANT MPS-VII VAL-354 AND TRP-611.  
 RX MEDLINE=94154731; PubMed=8111413;  
 RA Wu B.M., Sly W.S.,  
 RT "Mutational studies in a patient with the hydrops fetalis form of  
 RT mucopolysaccharidosis type VII.",  
 RL Hum. Mutat. 2:446-457(1993).  
 [10]  
 RP VARIANT MPS-VII CYS-382 AND VAL-619.  
 RX MEDLINE=91090114; PubMed=1702266;  
 RA Tomatsu S., Fukuda S., Sukeyama K., Ikeda Y., Yamada S., Yamada Y.,  
 RA Sasaki T., Okamoto H., Kuwahara T., Yamaguchi S., Kiman T.,  
 RA Shintaku H., Ishiki G., Oritani T.,  
 RT "Mucopolysaccharidosis type VII: characterization of mutations and  
 RT molecular heterogeneity.",  
 RL Am. J. Hum. Genet. 48:89-96(1991).  
 [11]  
 RP VARIANT MPS-VII CYS-627.  
 RX MEDLINE=93190983; PubMed=7680524;  
 RA Shipley J.W., Klinckenberg M., Wu B.M., Bachinsky D.R., Grubb J.H.,  
 RT "Mutational analysis of a patient with mucopolysaccharidosis type  
 RT VII, and identification of pseudogenes.",  
 RL Am. J. Hum. Genet. 52:517-526(1993).  
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE DEGRADATION OF DERMATAN  
 CC AND KERATAN SULFATES.  
 CC -1- CATALYTIC ACTIVITY: A beta-D-glucuronoside + H(2)O = an alcohol +  
 CC D-glucuronate.  
 CC -1- SUBUNIT: Homotetramer.  
 CC -1- SUBCELLULAR LOCATION: Lysosomal.

CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Long;  
 CC IsoId=PO8236-1; Sequence=Displayed;  
 CC Name=Short;  
 CC IsoId=PO8236-2; Sequence=VSP\_001799;  
 CC -1- PTM: GLYCOSYLATED WITH 3 TO 4 N-LINKED OLIGOSACCHARIDE CHAINS.  
 CC -1- DISEAS: Defects in GUSB are the cause of mucopolysaccharidosis  
 CC type VII (MPS-VII) [MIM:253220]; also known as Sly syndrome. This  
 CC is an autosomal recessive disorder characterized by excessive  
 CC dermatan and heparan sulfates in the urine and Hunter-like  
 CC features.  
 CC -1- SIMILARITY: Belongs to family 2 of glycosyl hydrolases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL, M15182; AAA52561.1; -;  
 CC EMBL, BC014142; AAH14142.1; -;  
 CC EMBL, M65002; AAA52622.1; -;  
 CC EMBL, M10618; AAA52621.1; -;  
 CC EMBL, 872462; AAD14101.1; -;  
 CC PIR, A26581; A26581.  
 CC PDB, 1BHG; 17-SEP-97.  
 CC Genew; HGNC:4696; GUSB.  
 CC MIM, 253220; -;  
 CC GO, GO:0004566; F:beta-glucuronidase activity; TAS.  
 CC GO, GO:0005975; P:carbohydrate metabolism; TAS.  
 CC GO, GO:0006027; P:glycosaminoglycan catabolism; TAS.  
 CC InterPro; IPR008979; Gal\_bind\_like.  
 CC InterPro; IPR006101; Glyco\_hydro\_2.  
 CC InterPro; IPR006102; Glyco\_hydro\_219.  
 CC InterPro; IPR006104; Glyco\_hydro\_25B.  
 CC InterPro; IPR006103; Glyco\_hydro\_2TMM.  
 CC Pfam; PF00703; Glyco\_hydro\_2; 1.  
 CC Pfam; PF02836; Glyco\_hydro\_2\_C; 1.  
 CC Pfam; PF02837; Glyco\_hydro\_2\_N; 1.  
 CC PRINTS; PR00132; GLHYDRASE2.  
 CC PROSITE; PS00719; GLYCOSYL\_HYDROL\_F2\_1; 1.  
 CC PROSITE; PS00608; GLYCOSYL\_HYDROL\_F2\_2; 1.  
 CC Hydrolase; Glycosidase; Lysosome; Glycoprotein; Signal;  
 CC Mucopolysaccharidosis; Disease mutation; 3D-structure;  
 CC KW Alternative splicing; Polymorphism.  
 CC FT SIGNAL 1 22  
 CC FT CHAIN 23 651  
 CC FT ACT\_SITE 451 451  
 CC FT CARBOHYD 173 173  
 CC FT CARBOHYD 272 272  
 CC FT CARBOHYD 420 420  
 CC FT CARBOHYD 631 631  
 CC FT VARSPLIC 505 355  
 CC FT VARIANT 216 216  
 CC FT VARIANT 354 354  
 CC FT VARIANT 382 382  
 CC FT VARIANT 611 611  
 CC FT VARIANT 619 619  
 CC FT VARIANT 627 627  
 CC FT VARIANT 649 649  
 CC FT STRAND 37 37  
 CC FT STRAND 43 43  
 CC  
 CC BETA-GLUCURONIDASE.  
 CC PROTON DONOR.  
 CC N-LINKED (GLCNAC. . .)  
 CC N-LINKED (GLCNAC. . .)  
 CC N-LINKED (GLCNAC. . .)  
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC Missing (in isoform Short).  
 CC /FTid=VSP\_001799.  
 CC R -> W (in MPS-VII).  
 CC /FTid=VAR\_003196.  
 CC A -> V (in MPS-VII).  
 CC /FTid=VAR\_003197.  
 CC R -> C (in MPS-VII).  
 CC /FTid=VAR\_003198.  
 CC R -> W (in MPS-VII).  
 CC /FTid=VAR\_003199.  
 CC A -> V (in MPS-VII).  
 CC /FTid=VAR\_003200.  
 CC W -> C (in MPS-VII).  
 CC /FTid=VAR\_003201.  
 CC P -> L (in dSNP:9530).  
 CC /FTid=VAR\_016179.

FT	STRAND	46	50
FT	HELIX	57	60
FT	TURN	61	62
FT	HELIX	63	65
FT	HELIX	68	71
FT	STRAND	75	76
PT	STRAND	79	79

Query Match 61.7%; Score 37; DB 1; Length 651;  
Best Local Similarity 58.3%; Pred. No. 32;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy	1	APPARVVMAVG	12
		:	
Db	2	ARGSAVWADLG	13

Search completed: September 1, 2004, 00:10:45  
Job time : 5.24667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2004, 23:56:54 ; Search time 22.24 Seconds

(without alignments)  
170.244 Million cell updates/sec

Title: US-09-508-095-22

Sequence: 1 ARRARYMAVAVG 12

Scoring table: BLOSUM62

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SPTRMBL\_25:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phase:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_virus:\*  
17: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	93.3	711	4	Q8TCD2
2	93.3	711	4	Q8TCD2
3	93.3	711	4	Q8TCD2
4	93.3	711	4	Q8TCD2
5	93.3	711	4	Q8TCD2
6	93.3	711	4	Q8TCD2
7	93.3	711	4	Q8TCD2
8	93.3	711	4	Q8TCD2
9	93.3	711	4	Q8TCD2
10	93.3	711	4	Q8TCD2
11	93.3	711	4	Q8TCD2
12	93.3	711	4	Q8TCD2
13	93.3	711	4	Q8TCD2
14	93.3	711	4	Q8TCD2
15	93.3	711	4	Q8TCD2
16	93.3	711	4	Q8TCD2

17	40	66.7	939	5	Q9VB35	Q9VB35 drosophila
18	39	65.0	88	10	Q84Z79	Q84Z79 oryza sativ
19	39	65.0	146	11	Q9DAB9	Q9DAB9 mus musc
20	39	65.0	477	16	Q7W83	Q7W83 bordetella
21	39	65.0	477	16	Q7W87	Q7W87 bordetella
22	39	65.0	493	17	Q8U3M8	Q8U3M8 pyrococcus
23	39	65.0	662	10	Q8W430	Q8W430 titicum de
24	38	63.3	18	6	Q9TR38	Q9TR38 bos taurus
25	38	63.3	295	16	Q829B8	Q829B8 streptomyce
26	38	63.3	307	16	Q8YF12	Q8YF12 bruceella me
27	38	63.3	307	16	Q8G2S8	Q8G2S8 bruceella me
28	38	63.3	322	16	Q89QD0	Q89QD0 bradyrhizob
29	38	63.3	342	10	Q82517	Q82517 saccharum h
30	38	63.3	431	16	Q910G9	Q910G9 streptomyce
31	38	63.3	438	16	Q834A4	Q834A4 enterococcu
32	38	63.3	467	16	Q829M0	Q829M0 streptomyce
33	38	63.3	490	10	Q7XTK7	Q7XTK7 zea mays (m
34	38	63.3	509	10	Q417S5	Q417S5 zea mays (m
35	38	63.3	567	10	Q65342	Q65342 saccharum o
36	38	63.3	567	10	Q65341	Q65341 saccharum r
37	38	63.3	640	10	Q7XZ55	Q7XZ55 saccharum h
38	38	63.3	655	10	Q9LKI9	Q9LKI9 oryza sativ
39	38	63.3	986	16	Q9A4S7	Q9A4S7 caulobacter
40	37	61.7	93	2	Q93T31	Q93T31 azospirillum
41	37	61.7	108	17	Q9YAA3	Q9YAA3 aeropyrum p
42	37	61.7	212	10	Q7XHM4	Q7XHM4 oryza sativ
43	37	61.7	256	16	Q9KZAS	Q9KZAS streptomyce
44	37	61.7	302	16	Q82N27	Q82N27 streptomyce
45	37	61.7	338	9	Q64065	Q64065 bacteriophag

## ALIGNMENTS

RESULT 1  
ID Q8TCD2 PRELIMINARY: PRT: 711 AA.

AC Q8TCD2; 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC EMBL: BC022347; AAH2347.1; -!  
DR GO: GO:0005576: C:extracellular; IEA.  
DR GO: GO:0008199: F:ferric iron binding; IEA.  
DR GO: GO:0006879: P:iron ion homeostasis; IEA.  
DR GO: GO:0006826: P:iron ion transport; IEA.  
DR GO: GO:0006810: P:iron ion transport; IEA.  
DR InterPro: IPR001156; Transferrin.  
DR Pfam: PF00405; transferrin; 2.  
DR PRINTS: PR00422; TRANSFERRIN.  
DR SMART: SM00094; TR\_FER; 2.  
DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE: PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE: PS00207; TRANSFERRIN\_3; 2.  
KW Glycoprotein; Iron transport; Metal-binding; Transport.  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7BE097C45FAF CRC64;  
SQ

Query Match 93.3%; Score 56; DB 4; Length 711;  
Best Local Similarity 91.7%; Pred. No. 0.24;

```

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ARARVWMAVG 12
Db 360 ARARVWCAVG 371

RESULT 2
Q81ZH6 PRELIMINARY; PRT; 711 AA.
ID Q81ZH6;
AC Q81ZH6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Lactoferrin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaplan J.B., Fine D.H.;
RT "Characterization of an amino acid polymorphism in the antibacterial
RT domain of human lactoferrin."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137470; AN11304.1;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006826; P:iron ion transport; IEA.
DR InterPro; IPR01156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PRO0422; TRANSFERRIN.
DR SMART; SMO0094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
SQ SEQUENCE 711 AA; 78396 MW; 547AB9423C27C667 CRC64;

Query Match 93.3%; Score 56; DB 4; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.24;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ARARVWMAVG 12
Db 360 ARARVWCAVG 371

RESULT 3
Q81U92 PRELIMINARY; PRT; 711 AA.
ID Q81U92;
AC Q81U92;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Lactotransferrin precursor (lactoferrin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y.-Q., Zhang Y., Zheng Y.-M.;
RT "Homo sapiens lactotransferrin Gene: CDNA Cloning and Sequence
RT Analysis."
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Baskar Singh S., Saravanan K., Paramasivan M., Srinivasan A.,
RA Singh T.P.;

```

```

RT "Homo sapiens lactoferrin (HLF) mRNA."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY165046; AN63998.1;
DR EMBL; AY178998; AN75578.2;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006826; P:iron ion transport; IEA.
DR InterPro; IPR01156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PRO0422; TRANSFERRIN.
DR SMART; SMO0094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
DR Glycoprotein; Iron transport; Metal-binding; Transport.
KW NON TER
FT SIGNAL
FT CHAIN
SQ SEQUENCE 711 AA; 78382 MW; 547BF642C9267E67 CRC64;

Query Match 93.3%; Score 56; DB 4; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.24;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ARARVWMAVG 12
Db 360 ARARVWCAVG 371

RESULT 4
Q8WN8 PRELIMINARY; PRT; 704 AA.
ID Q8WN8;
AC Q8WN8;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Lactoferrin (Pigment).
GN LTF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang S.-R., Lin T.-Y., Wang C.-N.;
RT "Isolation and expression of porcine milk lactoferrin."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC EMBL; L77887; AL40161.1;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006826; P:iron ion transport; IEA.
DR InterPro; IPR01156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PRO0422; TRANSFERRIN.
DR SMART; SMO0094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 1.
KW Glycoprotein; Iron transport; Metal-binding; Transport.
KW NON TER
FT SIGNAL
FT CHAIN
SQ SEQUENCE 704 AA; 77681 MW; 64EE769F503CC32 CRC64;

Query Match 81.7%; Score 49; DB 6; Length 704;
Best Local Similarity 75.0%; Pred. No. 3.8;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 ARRAVWMAVG 12  
 DB 355 AROAKVWCAVG 366

## RESULT 5

07YS20 PRELIMINARY; PRT; 704 AA.  
 ID Q7YS20  
 AC Q7YS20  
 DT 01-OCT-2003 (TREMBlrel. 25, Created)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Lactoferrin.  
 GN L.P.F.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OX NCBI\_TaxID=9623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Pecorini C., Fogher C., Baldi A.;  
 RT "The nucleotide sequence of porcine lactoferrin cDNA."  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY306198; AAP70487.1; -  
 SQ SEQUENCE 704 AA; 77522 MW; AACBE1767B56B56A CRC64;

Query Match 81.7%; Score 49; DB 6; Length 704;  
 Best Local Similarity 75.0%; Pred. No. 3.8;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ARRAVWMAVG 12  
 DB 355 AROAKVWCAVG 366

## RESULT 6

09SM57 PRELIMINARY; PRT; 112 AA.  
 ID 09SM57  
 AC 09SM57  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Lactoferrin (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE OF 1-40 FROM N.A.  
 RA Li G., Zhang Y., Li N.;  
 RT "Analysis of the 5'-region of the bovine lactoferrin gene using PCR-SSCP."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY036583; AAK66816.1; -  
 DR EMBL; AY036581; AAK66816.1; JOINED.  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0008199; F:ferric iron binding; IEA.  
 DR GO; GO:0006879; P:ferric iron binding; IEA.  
 DR GO; GO:0006826; P:iron ion homeostasis; IEA.  
 DR InterPro; IPR001156; Transferin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 FT NON TER 1  
 FT NON TER 112  
 SQ SEQUENCE 112 AA; 12491 MW; CFE0C01CB3F953FBA CRC64;

Query Match 75.0%; Score 45; DB 6; Length 112;  
 Best Local Similarity 75.0%; Pred. No. 3.1;  
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ARRAVWMAVG 12

DB 66 ARYTRVWCAVG 77

## RESULT 7

08MI10 PRELIMINARY; PRT; 355 AA.  
 ID 08MI10  
 AC 08MI10  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Lactotransferrin (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=91160550; PubMed=2001696;  
 RA Pierce A., Colavizza D., Benalissa M., Maes P., Tartar A.,  
 RA Montreuil J., Spix G.;  
 RT "Molecular cloning and sequence analysis of bovine lactotransferrin."  
 RL Eur. J. Biochem. 196;177-184 (1991).  
 DR EMBL; X17066; CAA34912.1; -  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0008199; F:ferric iron binding; IEA.  
 DR GO; GO:0006879; P:ferric iron homeostasis; IEA.  
 DR GO; GO:0006826; P:iron ion transport; IEA.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SMO0094; TR\_FER; 1.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 1.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 1.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 1.  
 FT NON TER 1  
 FT NON TER 355  
 SQ SEQUENCE 355 AA; 38861 MW; 92CF0C274EB6C893 CRC64;

Query Match 75.0%; Score 45; DB 6; Length 355;  
 Best Local Similarity 75.0%; Pred. No. 9.6;  
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ARRAVWMAVG 12  
 DB 6 ARYTRVWCAVG 17

## RESULT 8

059702 PRELIMINARY; PRT; 491 AA.  
 ID 059702  
 AC 059702  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE P-Hydroxybenzaldehyde dehydrogenase.  
 GN PCH.  
 OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=303;  
 RN [1]  
 RP PARTIAL SEQUENCE FROM N.A.  
 RC STRAIN=NCIMB 9856;  
 RA MEDLINE=95014081; PubMed=7929007;  
 RA Kim J., Fuller J.H., Cecchini G., McIntire W.S.;  
 RT "Cloning, sequencing, and expression of the structural genes for the  
 RT cytochrome and flavoprotein subunits of p-cresol methylhydroxylase  
 RL from two strains of Pseudomonas putida."  
 RN J. Bacteriol. 176:6349-6361 (1994).  
 RN [2]  
 SQ SEQUENCE FROM N.A.

```

RC STRAIN=NCIMB 9866;
RX MEDLINE=2002969; PubMed=10565539;
RA Cronin C.N., Kim J.-H., Fuller J., Zhang X.-P., McIntire W.S.;
RT "Organization and sequences of p-hydroxybenzaldehyde dehydrogenase and
other plasmid-encoded genes for early enzymes of the p-cresol
degradative pathway in Pseudomonas putida NCIMB 9866 and 9869."
RT DNA Seq. 10:7-17(1999).
RL (3)
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 9866;
RA Cronin C.N.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
DR EMBL; U96338; AAAY5634.2; -.
DR PIR; T46884; T46884.
DR HSSP; P51977; 1BX5.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0014691; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldenhyde_dehydr.
DR Pfam; PF00171; aldehyd_1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU_1.
DR Oxidoreductase; Plasmid.
KW SEQUENCE 491 AA; 53156 MW; 70C45CE421266344 CRC64;

Query Match 71.7%; Score 43; DB 2; Length 491;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ARARVWMAVG 12
Db 58 ARQAGVMAACG 69

RESULT 9
Q8CBA0 PRELIMINARY; PRT; 707 AA.
ID Q8CBA0;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE Lactoferrin.
GN LTF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RL EMBL; AK036491; BAC29450.1; -.
DR MGD; MGI:96837; Ltf.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0006879; P:iron ion homeostasis; IEA.
DR GO; GO:0006826; P:iron ion transport; IEA.
DR InterPro; IPR001156; transferrin.
DR Pfam; PF00405; transferrin_2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER_2.
DR PROSITE; PS00205; TRANSFERRIN_1; 1.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW SEQUENCE 707 AA; 77837 MW; E1B32F5FD8748A0F CRC64;

Query Match 71.7%; Score 43; DB 11; Length 707;

```

```

Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARARVWMAVG 12
Db 358 ASKARVWCAVG 369

RESULT 10
Q82HU3 PRELIMINARY; PRT; 254 AA.
ID Q82HU3;
AC Q82HU3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE Putative short-chain dehydrogenase.
GN SAV3415.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary
metabolites."
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608305; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005035; BAC1127.1; -.
DR GO; GO:0014691; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short_1.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Complete proteome.
KW SEQUENCE 254 AA; 26802 MW; 66CB807E046409F0 CRC64;

Query Match 68.3%; Score 41; DB 16; Length 254;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ARARVWMA 10
Db 33 ARARVWMA 42

RESULT 11
Q93VW3 PRELIMINARY; PRT; 108 AA.
ID Q93VW3;
AC Q93VW3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE B1148D12.15 protein (P0454H12.9 protein).
GN B1148D12.15 OR P0454H12.9.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;

```



```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
   clone:B148D12."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
   clone:P0454H12."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003411; BAB64830.1; -
DR EMBL; AP003255; BAB62577.1; -
DR Gramene; O93YW3; -
SQ SEQUENCE 108 AA; 11462 MW; 81E613543D906340 CRC64;

Query Match
Best Local Similarity 66.7%; Score 40; DB 10; Length 108;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12
DB 91 ABAEAVMAVG 102

RESULT 12
Q89C14 PRELIMINARY; PRT; 187 AA.
AC Q89C14;
AT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Transcriptional regulatory protein.
GN BLR7984.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Matsumoto M., Irida K., Irida M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Teurroka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
   Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005964; BAC53249.1; -
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR01647; HTH_Tetr.
DR Pfam; PF00440; Tetr_1.
DR PRINTS; PR00455; HTHTEIR.
DR Complete Proteome.
SQ SEQUENCE 187 AA; 20135 MW; B60CF64F900A222 CRC64;

Query Match
Best Local Similarity 66.7%; Score 40; DB 16; Length 187;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRARVMAVG 12
DB 147 RRARAFIAIG 157

RESULT 13
Q866X7 PRELIMINARY; PRT; 328 AA.

```

```

AC Q866X7;
AT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Carbonic anhydrase-related XI protein.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RP SEQUENCE FROM N.A.
RA Chen Y., Huang C.-H.;
RT "Molecular identification of carbonic anhydrases (CA) and CA-related
   (CAR) genes."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY075023; AAL78171.1; -
DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0008730; F:one-carbon compound metabolism; IEA.
DR InterPro; IPR01148; Euk_Coand.
DR Pfam; PF00194; carb_anhydrase; 1.
DR Prodom; PD000865; Euk_Coand; 1.
SQ SEQUENCE 328 AA; 36144 MW; 29C7249B63D5006 CRC64;

Query Match
Best Local Similarity 66.7%; Score 40; DB 6; Length 328;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARRAVWMAVG 12
DB 9 APRALVMAVG 20

RESULT 14
Q7U292 PRELIMINARY; PRT; 449 AA.
AC Q7U292;
AT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Possible cytochrome P450 135A1 CYP135A1 (EC 1.14.-).
GN CYP135A1 OR MB0334C.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1763;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12786972;
RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Durhoy S., Gordon S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248335; CAD93198.1; -
DR Oxidoreductase; Complete Proteome.
SQ SEQUENCE 449 AA; 50036 MW; F68D71CCB18F7BE4 CRC64;

Query Match
Best Local Similarity 70.0%; Score 40; DB 16; Length 449;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RARVWMAVG 12
DB 285 QSOVWMAVG 294

RESULT 15
Q98923 PRELIMINARY; PRT; 504 AA.

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2004, 23:54:24 ; Search time 24.0667 Seconds

(without alignments)  
117.402 Million cell updates/sec

Title: US-09-508-095-8

Perfect score: 50

Sequence: 1 EVAAARARVW 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A\_Geneseq\_290a04:\*  
1: Geneseq180a:\*  
2: Geneseq190a:\*  
3: Geneseq200a:\*  
4: Geneseq2001a:\*  
5: Geneseq2002a:\*  
6: Geneseq2003a:\*  
7: Geneseq2003b:\*  
8: Geneseq2004a:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	10	2	AAW93872
2	41	82.0	980	2	AAW93872 Bifidobac
3	39.5	79.0	687	2	AAW93872 Recombina
4	39.5	79.0	687	2	AAW93872 Mutant hu
5	39.5	79.0	688	2	AAW93872 Mutant hu
6	39.5	79.0	688	2	AAW93872 Mutant hu
7	39.5	79.0	688	2	AAW93872 Mutant hu
8	39.5	79.0	688	2	AAW93872 Mutant hu
9	39.5	79.0	688	2	AAW93872 Mutant hu
10	39.5	79.0	688	2	AAW93872 Mutant hu
11	39.5	79.0	688	2	AAW93872 Mutant hu
12	39.5	79.0	688	2	AAW93872 Mutant hu
13	39.5	79.0	688	2	AAW93872 Mutant hu
14	39.5	79.0	688	2	AAW93872 Mutant hu
15	39.5	79.0	688	2	AAW93872 Mutant hu
16	39.5	79.0	688	2	AAW93872 Mutant hu
17	39.5	79.0	688	2	AAW93872 Mutant hu
18	39.5	79.0	688	2	AAW93872 Mutant hu
19	39.5	79.0	688	2	AAW93872 Mutant hu
20	39.5	79.0	688	2	AAW93872 Mutant hu
21	39.5	79.0	688	2	AAW93872 Mutant hu
22	39.5	79.0	688	2	AAW93872 Mutant hu
23	39.5	79.0	688	2	AAW93872 Mutant hu
24	39.5	79.0	688	2	AAW93872 Mutant hu
25	39.5	79.0	688	2	AAW93872 Mutant hu

26	39.5	79.0	709	2	AAW53879
27	39.5	79.0	709	3	AAW77577
28	39.5	79.0	709	3	AAW03830
29	39.5	79.0	709	4	AAW36842
30	39.5	79.0	709	7	AAW76659
31	39.5	79.0	711	2	AAW08033
32	39.5	79.0	711	2	AAW43653
33	39.5	79.0	711	2	AAW09342
34	39.5	79.0	711	2	AAW57317
35	39.5	79.0	711	2	AAW53880
36	39.5	79.0	711	2	AAW86021
37	39.5	79.0	711	3	AAW77578
38	39.5	79.0	711	3	AAW03831
39	39.5	79.0	711	3	AAW08182
40	39.5	79.0	711	4	AAW36843
41	39.5	79.0	711	4	AAW02341
42	39.5	79.0	711	4	AAW64828
43	39.5	79.0	711	4	AAW64661
44	39.5	79.0	711	4	AAW64791
45	39.5	79.0	711	6	AAW61310

## ALIGNMENTS

RESULT 1	AAW93872	standard; peptide; 10 AA.
ID	AAW93872	standard; peptide; 10 AA.
XX	AAW93872	
AC	AAW93872	
DT	27-AUG-2003 (revised)	
DT	25-JUN-1999 (first entry)	
DE	Bifidobacterium bifidus stimulating peptide 8.	
XX		
KW	Bifidogenic peptide; protease; treatment; microbe-related disease; bacteria; fungi; yeast; protozoa; virus; mycoplasma; filaria; plasmodia; infection; inflammation; microbially induced tumour; degenerative disorder; diarrhoea; colic; oral microflora; intestinal microflora; caries; vaginal microflora.	
OS	Bifidobacterium bifidum.	
XX		
PN	W09914231-A2.	
PD	25-MAR-1999.	
XX		
PF	16-SEP-1998; 98WO-EP005899.	
XX		
PR	16-SEP-1997; 97DE-01040604.	
PR	11-FEB-1998; 98DE-01005385.	
XX		
PA	(FORS/) FORSMANN W.	
XX		
PI	Forsmann W, Zucht H, Liepke C;	
XX		
DR	WPI, 1999-244022/20.	
XX		
PT	Milk-derived peptides that stimulate Bifidobacterium bifidus.	
XX		
PS	Claim 2; Page 3; 25p; German.	
XX		
CC	This invention describes milk-derived bifidogenic peptides and their active derivatives or fragments, and combinations of them produced by chemical coupling. Such are produced from bovine or human milk by treatment for 2 hr with proteases, then centrifuging to remove fat and acidifying to pH 2 to precipitate proteins. The solution phase is then subjected to reverse-phase high-performance liquid chromatography (HPLC) and cation-exchange HPLC, the fractions adjusted to salt content below 25 mM (by dialysis or reverse-phase HPLC) and tested for activity by growing Bifidobacterium bifidus and Escherichia coli in presence of the fractions. Those fractions for which (Bw-B0)-(Ew-E0) is at least 0.15 are	

CC selected where BW = germ count after 16 hr culture of *B. bifidus* in 50%  
 CC Elliker broth containing peptide at 0.2 mg/ml. B0 = germ count under  
 CC similar conditions in a peptide-free control. Ew = germ count after 16 hr  
 CC culture of *E. coli* in 3 g/l tryptic broth containing peptide at 0.2  
 CC mg/ml. B0 = germ count under similar conditions in a peptide-free  
 CC control. The peptides AAW93865-W93888 are used to treat microbe-related  
 CC diseases caused by bacteria, fungi, yeast, protozoa, viruses, mycoplasma,  
 CC filaria and plasmodia, e.g. infections, inflammation, microbially induced  
 CC tumours or degenerative disorders, diarrhoea, colic, abnormalities in  
 CC oral, intestinal or vaginal microflora, or carries. (Updated on 27-AUG-  
 CC 2003 to correct OS field.)  
 XX

SO Sequence 10 AA;

Query Match 100.0%; Score 50; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVARARVVM 10  
 |||||  
 Db 1 EVARARVVM 10

RESULT 2  
 AAB62027

ID AAB62027 standard; protein; 980 AA.

XX AAB62027;

DT 14-MAY-2001 (first entry)

DE Recombinant *P. furiosus* helicase 4.

KM MCM; minichromosome maintenance protein; archaeal polypeptide; PCNA;  
 KM RFC-P38; RFC-P55; RPA; CD66; FEN-1; dTTPase; ligase; helicase dna2; PCR;  
 KM helicase 4; nucleic acid amplification; polymerase chain reaction.

OS Pyrococcus furiosus.

XX WO200109347-A2.

XX 08-FEB-2001.

XX 28-JUL-2000; 2000WO-US020532.

XX 30-JUL-1999; 99US-0146580P.

XX (STRA-) STRATAGENE.

XX Hogrefe HH, Cline JM, Hansen CJ, Borns MC;

XX WPI; 2001-182959/18.

XX N-PSDB; AAF57031.

PT Composition for improving nucleic acid polymerase reactions, useful e.g.  
 PT in synthesis or amplification, contains at least one archaeal accessory  
 PT protein.

XX Claim 182; Fig 31; 147bp; English.

XX The invention provides a composition (A) for enhancing nucleic acid  
 CC polymerase reactions that comprises an archaeal MCM (minichromosome  
 CC maintenance protein) and at least one of the archaeal polypeptides (PCNA,  
 CC RFC-P38 or -P55, RPA, CDC6, FEN-1, dTTPase, ligase, helicase dna2, or  
 CC helicases 2-8). (A) And similar compositions containing different  
 CC combinations of accessory proteins, are used to improve performance of  
 CC synthesis, amplification, mutagenizing, labeling and detecting reactions,  
 CC e.g. for gene characterization, cloning, detection of allelic variants,  
 CC diagnosis and screening for disease, particularly where done by  
 CC polymerase chain reaction (PCR). Some of the proteins also stabilize  
 CC duplexes during polymerase reactions or improve exonuclease reactions,  
 CC for example RPA also improves specificity of nucleic acid/protein  
 CC interaction and PCNA improves polymerase-mediated repair processes and

CC hybridization reactions. Nucleic acids encoding the archaeal polypeptides  
 CC are used for recombinant production of proteins, and fragments of the  
 CC nucleic acid as probes and primers for screening related sequences. The  
 CC accessory proteins increase accuracy and efficiency of polymerase  
 CC reactions, allow use of lower denaturation and extension temperatures  
 CC (possibly isothermal processing), and improve synthesis of long targets.  
 CC The present sequence represents a *P. furiosus* recombinant helicase 4  
 XX

SO Sequence 980 AA;

Query Match 82.0%; Score 41; DB 4; Length 980;  
 Best Local Similarity 80.0%; Pred. No. 52;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVARARVVM 10  
 |||||  
 Db 298 EVARARVVM 307

RESULT 3  
 AAW71183

ID AAW71183 standard; protein; 687 AA.

XX AAW71183;

DT 27-OCT-1998 (first entry)

DE Mutant human lactoferrin protein designated hLF-5N.

XX Human; lactoferrin protein; variant; anti-inflammatory; anti-viral;  
 XX anti-infective; coagulant; complement activation; inhibition;  
 XX LPS mediated activation; myeloperoxidase; growth promotion;  
 XX intestinal epithelial cell; hydroxyl-radical formation;  
 XX intestinal iron uptake; excretion.

XX Synthetic.

OS Homo sapiens.

XX WO9833509-A2.

XX 06-AUG-1998.

XX 02-FEB-1998; 98WO-IB000441.

XX 03-FEB-1997; 97US-0036659P.

XX 02-FEB-1998; 98US-00017043.

XX (PHAR-) PHARMING BV.

XX Nuijens J, Van Berkel PHC;

XX WPI; 1998-437164/37.

PT Compositions containing human lactoferrin and variants - are used for  
 PT treating e.g. anaemia, iron-storage disease, inflammation, tumours,  
 PT rheumatoid arthritis, ulcerative colitis or infections.

XX Claim 2; Page; 70pp; English.

XX AAW71180-83 represent N-terminally truncated human lactoferrin proteins.  
 CC The lactoferrin variants bind heparin with lower affinity than natural  
 CC lactoferrin. The lactoferrin variants exhibit biological activities such  
 CC as anti-inflammatory, anti-viral and anti-infective activities as well as  
 CC a pro- and anti-coagulant effects, modulation of complement activation,  
 CC inhibition of LPS mediated activation of neutrophils, inhibition of  
 CC myeloperoxidase, regulation of transcription, growth promotion of intestinal  
 CC epithelial cells, inhibition of hydroxyl-radical formation, and a role in  
 CC intestinal iron uptake and excretion. note: this sequence does not appear  
 CC in the specification; it was created using information provided  
 XX

SO Sequence 687 AA;

Query Match 79.0%; Score 39.5; DB 2; Length 687;

Best Local Similarity 90.9%; Pred. No. 69;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10  
333 EVARRARVW 343

## RESULT 4

ID AAG77911 standard; protein; 687 AA.

AC AAG77911;

DT 22-JAN-2002 (first entry)

DE Human lactoferrin variant hLF-SN.

XX Human; lactoferrin; hLF; infectious disease; inflammatory disease;

KW excess of heparin; gastroenteritis; inflammatory bowel disease; sepsis;

KM anaemia; myelopoiesis; reperfusion injury; cytokine release; proteoglycan;

XX hLF-SN; mutant; murein.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 23..26 /label=Cationic\_domain

XX WO200172322-A2.

PD 04-OCT-2001.

PF 27-MAR-2001; 2001WO-NL000253.

XX 27-MAR-2000; 2000EP-00201110.

PR 27-MAR-2000; 2000US-0193352P.

XX (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.

PI Van Bree JBMW, Nuijens JH;

XX WPI; 2001-648424/74.

DR Use of lactoferrin for treatment of infectious diseases, inflammatory

XX diseases and excess of heparin.

PT Claim 10; Page; 49pp; English.

XX The sequence represents the human lactoferrin (hLF) protein variant hLF-

CC 5N. The invention relates to novel methods of treatment using high doses

CC of lactoferrin. The methods of the invention are useful for the treatment

CC or prophylaxis of infectious diseases, inflammatory diseases and excess

CC of heparin e.g. gastroenteritis, inflammatory bowel disease, sepsis,

CC anaemia, myelopoiesis, reducing reperfusion injury, cytokine release and

CC proteoglycan-mediated entry of virus into cells. The advantage of the

CC method is that the patient is substantially free of side effect responses

CC to administration of lactoferrin. Therefore large doses of lactoferrin

CC can be administered. Note: The present sequence is not shown in the

CC specification but is derived from human wild-type lactoferrin sequence

CC given in the sequence listing of the specification

CC SQ Sequence 687 AA;

Query Match

Best Local Similarity 79.0%; Score 39.5; DB 4; Length 687;

Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10

333 EVARRARVW 343

RESULT 5  
ID AAW71182 standard; protein; 688 AA.

AC AAW71182;

DT 27-OCT-1998 (first entry)

DE Mutant human lactoferrin protein designated hLF-4N.

XX Human; lactoferrin protein; variant; anti-inflammatory; anti-viral;

KW anti-infective; coagulant; complement activation; inhibition;

KM LPS mediated activation; myelopoiesis; growth promotion;

XX intestinal epithelial cell; hydroxyl-radical formation;

XX intestinal iron uptake; excretion.

OS Synthetic.

XX Homo sapiens.

XX WO9833509-A2.

PD 06-AUG-1998.

PF 02-FEB-1998; 98WO-IB000441.

XX 03-FEB-1997; 97US-0036859P.

PR 02-FEB-1998; 98US-00017043.

XX (PHAR-) PHARMING BV.

XX Nuijens J, Van Berkel PHC;

XX WPI; 1998-437164/37.

XX Compositions containing human lactoferrin and variants - are used for

XX treating e.g. anaemia, iron-storage disease, inflammation, tumours,

XX rheumatoid arthritis, ulcerative colitis or infections.

XX Claim 2; Page; 70pp; English.

XX AAW71180-83 represent N-terminally truncated human lactoferrin proteins.

XX The lactoferrin variants bind heparin with lower affinity than natural

XX lactoferrin. The lactoferrin variants exhibit biological activities such

XX as anti-inflammatory, anti-viral and anti-infective activities as well as

XX a pro- and anti-coagulant effects, modulation of complement activation,

XX inhibition of LPS mediated activation of neutrophils, inhibition of

XX myelopoiesis, regulation of transcription, growth promotion of intestinal

XX epithelial cells, inhibition of hydroxyl-radical formation, and a role in

XX intestinal iron uptake and excretion. note: this sequence does not appear

XX in the specification; it was created using information provided

XX SQ Sequence 688 AA;

Query Match

Best Local Similarity 79.0%; Score 39.5; DB 2; Length 688;

Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10

334 EVARRARVW 344

RESULT 6

ID AAG77910 standard; protein; 688 AA.

AC AAG77910;

DT 22-JAN-2002 (first entry)

DE Human lactoferrin variant hLF-4N.

XX Human; lactoferrin; hLF; infectious disease; inflammatory disease;

KW excess of heparin; gastroenteritis; inflammatory bowel disease; sepsis;  
 KW anaemia; myelopoiesis; reperfusion injury; cytokine release; proteoglycan;  
 KW hLF-4N; mutant; mutein.  
 OS Homo sapiens.  
 XX  
 XX  
 FT Key Location/Qualifiers  
 FT Domain 24..27  
 FT /label= Cationic\_domain  
 XX  
 XX WO200172322-A2.  
 PN  
 XX  
 PD 04-OCT-2001.  
 XX  
 XX 27-MAR-2001; 2001WO-NL000253.  
 PF  
 XX 27-MAR-2000; 2000EP-00201110.  
 PR  
 XX 27-MAR-2000; 2000US-0193352P.  
 XX  
 XX (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.  
 PA  
 XX  
 PI Van Bree JEMM, Nuijens JH;  
 DR WPI; 2001-648424/74.  
 XX  
 XX  
 PT Use of lactoferrin for treatment of infectious diseases, inflammatory  
 PT diseases and excess of heparin.  
 PS  
 PS Claim 10; Page; 49pp; English.  
 XX  
 XX The sequence represents the human lactoferrin (hLF) protein variant hLF-  
 CC 4N. The invention relates to novel methods of treatment using high doses  
 CC of lactoferrin. The methods of the invention are useful for the treatment  
 CC of propylaxis of infectious diseases; inflammatory diseases and excess  
 CC of heparin e.g. gastroenteritis, inflammatory bowel disease, sepsis,  
 CC anaemia, myelopoiesis, reducing reperfusion injury, cytokine release and  
 CC proteoglycan-mediated entry of virus into cells. The advantage of the  
 CC method is that the patient is substantially free of side effect responses  
 CC to administration of lactoferrin. Therefore large doses of lactoferrin  
 CC can be administered. Note: The present sequence is not shown in the  
 CC specification but is derived from human wild-type lactoferrin sequence  
 CC given in the sequence listing of the specification  
 CC  
 SQ Sequence 688 AA;  
 QY  
 DB 1 EVAA-PARVW 10  
 334 EVAAARRARVW 344  
 RESULT 7  
 ID AAW71181 standard; protein; 689 AA.  
 XX  
 XX AAW71181;  
 AC  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE Mutant human lactoferrin protein designated hLF-3N.  
 XX  
 XX Human; lactoferrin protein; variant; anti-inflammatory; anti-viral;  
 KW anti-infective; coagulant; complement activation; inhibition;  
 KW LPS mediated activation; myelopoiesis; growth promotion;  
 KW intestinal epithelial cell; hydroxyl-radical formation;  
 KW intestinal iron uptake; excretion.  
 XX  
 XX Synthetic.  
 OS Homo sapiens.  
 XX

PN WO9833509-A2.  
 XX  
 PD 06-AUG-1998.  
 XX  
 XX 02-FEB-1998; 98WO-IB000441.  
 PF  
 XX 03-FEB-1997; 97US-0036859P.  
 PR  
 XX 02-FEB-1998; 98US-00017043.  
 XX  
 XX (PHAR-) PHARMING BV.  
 PA  
 XX  
 PI Nuijens J, Van Berckel PHC;  
 DR WPI; 1998-437164/37.  
 XX  
 XX Compositions containing human lactoferrin and variants - are used for  
 PT treating e.g. anaemia, iron-storage disease, inflammation, tumours,  
 PT rheumatoid arthritis, ulcerative colitis or infections.  
 PS  
 PS Claim 2; Page; 70pp; English.  
 XX  
 XX AAW71180-83 represent N-terminally truncated human lactoferrin proteins.  
 CC The lactoferrin variants bind heparin with lower affinity than natural  
 CC lactoferrin. The lactoferrin variants exhibit biological activities such  
 CC as anti-inflammatory, anti-viral and anti-infective activities as well as  
 CC a pro- and anti-coagulant effects, modulation of complement activation,  
 CC inhibition of LPS mediated activation of neutrophils, inhibition of  
 CC myelopoiesis, regulation of transcription, growth promotion of intestinal  
 CC epithelial cells, inhibition of hydroxyl-radical formation, and a role in  
 CC intestinal iron uptake and excretion. Note: This sequence does not appear  
 CC in the specification; it was created using information provided  
 CC  
 SQ Sequence 689 AA;  
 QY  
 DB 1 EVAA-PARVW 10  
 335 EVAAARRARVW 345  
 RESULT 8  
 ID AAG77909 standard; protein; 689 AA.  
 XX  
 XX AAG77909;  
 AC  
 DT 22-JAN-2002 (first entry)  
 XX  
 DE Human lactoferrin variant hLF-3N.  
 XX  
 XX Human; lactoferrin; hLF; infectious disease; inflammatory disease;  
 KW excess of heparin; gastroenteritis; inflammatory bowel disease; sepsis;  
 KW anaemia; myelopoiesis; reperfusion injury; cytokine release; proteoglycan;  
 KW mutant; mutein; hLF-3N; mutant; mutein.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 FT Key Location/Qualifiers  
 FT Domain 25..28  
 FT /label= Cationic\_domain  
 XX  
 XX WO200172322-A2.  
 PN  
 XX  
 PD 04-OCT-2001.  
 XX  
 XX 27-MAR-2001; 2001WO-NL000253.  
 PF  
 XX 27-MAR-2000; 2000EP-00201110.  
 PR  
 XX 27-MAR-2000; 2000US-0193352P.  
 XX

PA (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.  
 XX  
 PI Van Bree JBMV, Nuijens JH;  
 XX  
 DR WPI, 2001-648424/74.  
 XX  
 PT Use of lactoferrin for treatment of infectious diseases, inflammatory  
 PT diseases and excess of heparin.  
 XX  
 PS Claim 10; Page; 49pp; English.  
 CC  
 CC The sequence represents the human lactoferrin (hLF) protein variant hLF-  
 CC 3N. The invention relates to novel methods of treatment using high doses  
 CC of lactoferrin. The methods of the invention are useful for the treatment  
 CC or prophylaxis of infectious diseases, inflammatory diseases and excess  
 CC of heparin e.g. gastroenteritis, inflammatory bowel disease, sepsis,  
 CC anaemia, myelopoieses, reducing reperfusion injury, cytokine release and  
 CC proteoglycan-mediated entry of virus into cells. The advantage of the  
 CC method is that the patient is substantially free of side effect responses  
 CC to administration of lactoferrin. Therefore large doses of lactoferrin  
 CC can be administered. Note: The present sequence is not shown in the  
 CC specification but is derived from human wild-type lactoferrin sequence  
 CC given in the sequence listing of the specification  
 XX  
 SQ Sequence 689 AA;  
 Query Match 79.0%; Score 39.5; DB 4; Length 689;  
 Best Local Similarity 90.9%; Pred. No. 69;  
 Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 EVAA-RAAYVM 10  
 DB 335 EVAARRARAYVM 345  
 RESULT 9  
 AAW71180  
 ID AAW71180 standard; protein; 690 AA.  
 AC AAW71180;  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE Mutant human lactoferrin protein designated hLF-2N.  
 XX  
 KW Human; lactoferrin protein; variant; anti-inflammatory; anti-viral;  
 KW anti-infective; coagulant; complement activation; inhibition;  
 KW LPS mediated activation; myelopoiesis; growth promotion;  
 KW intestinal epithelial cell; hydroxyl-radical formation;  
 KW intestinal iron uptake; excretion.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9833509-A2.  
 XX  
 PD 06-AUG-1998.  
 XX  
 PF 02-FEB-1998; 98WO-IB000441.  
 XX  
 PR 03-FEB-1997; 97US-0036859P.  
 PR 02-FEB-1998; 98US-00017043.  
 XX  
 PA (PHAR-) PHARMING BV.  
 XX  
 PI Nuijens J, Van Berkel PHC;  
 XX  
 DR WPI, 1998-437164/37.  
 XX  
 PT Compositions containing human lactoferrin and variants - are used for  
 PT treating e.g. anaemia, iron-storage disease, inflammation, tumours,  
 PT rheumatoid arthritis, ulcerative colitis or infections.  
 XX

PS Claim 2; Page; 70pp; English.  
 XX  
 CC AAW71180-83 represent N-terminally truncated human lactoferrin proteins.  
 CC The lactoferrin variants bind heparin with lower affinity than natural  
 CC lactoferrin. The lactoferrin variants exhibit biological activities such  
 CC as anti-inflammatory, anti-viral and anti-infective activities as well as  
 CC a pro- and anti-coagulant effects, modulation of complement activation,  
 CC inhibition of LPS mediated activation of neutrophils, inhibition of  
 CC myelopoiesis, regulation of transcription, growth promotion of intestinal  
 CC epithelial cells, inhibition of hydroxyl-radical formation, and a role in  
 CC intestinal iron uptake and excretion. note: this sequence does not appear  
 CC in the specification; it was created using information provided  
 XX  
 SQ Sequence 690 AA;  
 Query Match 79.0%; Score 39.5; DB 2; Length 690;  
 Best Local Similarity 90.9%; Pred. No. 69;  
 Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 EVAA-RAAYVM 10  
 DB 336 EVAARRARAYVM 346  
 RESULT 10  
 AAG77908  
 ID AAG77908 standard; protein; 690 AA.  
 AC AAG77908;  
 DT 22-JAN-2002 (first entry)  
 XX  
 DE Human lactoferrin variant hLF-2N.  
 XX  
 KW Human; lactoferrin; hLF; infectious disease; inflammatory disease;  
 KW excess of heparin; gastroenteritis; inflammatory bowel disease; sepsis;  
 KW anaemia; myelopoiesis; reperfusion injury; cytokine release;  
 KW proteoglycan; hLF-2N; mutant; mutein.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 26..29  
 FT /label= Cationic\_domain  
 XX  
 PN WO200172322-A2.  
 XX  
 PD 04-OCT-2001.  
 XX  
 PF 27-MAR-2001; 2001WO-NL000253.  
 XX  
 PR 27-MAR-2000; 2000EP-00201110.  
 PR 27-MAR-2000; 2000US-0193352P.  
 XX  
 PA (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.  
 XX  
 PI Van Bree JBMV, Nuijens JH;  
 XX  
 DR WPI, 2001-648424/74.  
 XX  
 PT Use of lactoferrin for treatment of infectious diseases, inflammatory  
 PT diseases and excess of heparin.  
 XX  
 PS Claim 10; Page; 49pp; English.  
 CC  
 CC The sequence represents the human lactoferrin (hLF) protein variant hLF-  
 CC 2N. The invention relates to novel methods of treatment using high doses  
 CC of lactoferrin. The methods of the invention are useful for the treatment  
 CC or prophylaxis of infectious diseases, inflammatory diseases and excess  
 CC of heparin e.g. gastroenteritis, inflammatory bowel disease, sepsis,  
 CC anaemia, myelopoieses, reducing reperfusion injury, cytokine release and  
 CC proteoglycan-mediated entry of virus into cells. The advantage of the  
 CC method is that the patient is substantially free of side effect responses

CC to administration of lactoferrin. Therefore large doses of lactoferrin  
 CC can be administered. Note: The present sequence is not shown in the  
 CC specification but is derived from human wild-type lactoferrin sequence  
 CC given in the sequence listing of the specification  
 XX

SO Sequence 690 AA;

Query Match 79.0%; Score 39.5; DB 4; Length 690;  
 Best Local Similarity 90.9%; Pred. No. 69;  
 Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RAVVM 10  
 |||||  
 DB 336 EVAAARARVVM 346

RESULT 11  
 AEG80724  
 ID AEG80724 standard; protein; 690 AA.

AC AEG80724;

DT 29-NOV-2002 (first entry)

DE Codon optimised lactoferrin protein.

KM Transformed plant; heterologous transcription factor; transgenic plant;  
 seed protein; protein expression.

OS Homo sapiens.

PN WO200264750-A2.

PD 22-AUG-2002.

PF 14-FEB-2002; 2002MO-US004909.

PR 14-FEB-2001; 2001US-0269188P.

PR 14-FEB-2001; 2001US-0269199P.

PR 02-MAY-2001; 2001US-00847232.

PA (VENT-) VENTRIA BIOSCIENCE.

PI Huang N, Yang D;

DR WPI; 2002-657592/70.

DR N-PSDB; ABS66515.

PT Producing heterologous polypeptide in plant grain, by culturing  
 PT transformed plant to form a grain-producing transforming plant, and  
 PT recovering transgenic grains containing polypeptide from transformed  
 PT plant.

PS Example 15; Page 117; 230pp; English.

XX The invention describes a method of producing a heterologous polypeptide  
 CC (1) in a grain of a plant, comprising culturing a transformed plant (P1)  
 CC comprising a first chimeric gene, and optionally, at least one  
 CC heterologous transcription factor that is capable of enhancing the  
 CC expression of the chimeric gene, to form a grain producing transforming  
 CC plant (P2), and recovering transgenic grains containing (1) from P2. The  
 CC method is useful for producing heterologous polypeptide in a grain of a  
 CC plant. (1) is a non-plant storage, human or non-human animal, milk or  
 CC other than a milk polypeptide, antibodies, cytokines, lymphokines,  
 CC chemokines, hormones, growth factors, coagulation factors, anti-  
 CC infectives, or cytotoxins, or anti-inflammatory molecule or intestinal  
 CC trefoil factor (ITF) or its active fragment. Preferably, (1) is  
 CC lactoferrin, lysozyme, lactoferricin, ITF, epidermal growth factor (EGF),  
 CC keratinocyte growth factor (KGF), insulin-like growth factor I (IGF-I),  
 CC lactomedrin, kappa-casein, haptocorrin, lactoperoxidase, alpha-1-  
 CC antitrypsin, immunoglobulins, alpha-lactalbumin, beta-lactoglobulin,  
 CC alpha-casein, beta-casein, albumin, fibrinogen or protease inhibitor.  
 CC This is the amino acid sequence of a protein associated with method of

CC producing a transgenic plant  
 XX

SO Sequence 690 AA;

Query Match 79.0%; Score 39.5; DB 5; Length 690;  
 Best Local Similarity 90.9%; Pred. No. 69;  
 Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RAVVM 10  
 |||||  
 DB 336 EVAAARARVVM 346

RESULT 12  
 AAE27884  
 ID AAE27884 standard; protein; 690 AA.

AC AAE27884;

DT 27-DEC-2002 (first entry)

DE Human codon optimised lactoferrin.

KM Human; feed improvement; plant-derived feed; antibiotic; additive;  
 anti-microbial; poultry; lactoferrin; flour; malt.

OS Homo sapiens.

OS Synthetic.

PH Key Location/Qualifiers

FT Misc-difference 319..320  
 /note="Encoded by CTG TAC CTC"

PN WO200263975-A2.

PD 22-AUG-2002.

PF 14-FEB-2002; 2002MO-US004919.

PR 14-FEB-2001; 2001US-0269188P.

PR 02-MAY-2001; 2001US-00847232.

PA (VENT-) VENTRIA BIOSCIENCE.

PI Huang N, Rodriguez RL, Hagie PE;

DR WPI; 2002-682708/73.

DR N-PSDB; AAD45297.

PT Improved feed for production animals, comprising plant-derived feed  
 PT ingredients, and seed composition containing flour, extract, or malt from  
 PT mature monocot seeds and heterologous seed-produced anti-microbial  
 PT proteins.

PS Example 7; Page 148-150; 175pp; English.

XX The invention relates to improved feed for production animals, comprising  
 CC one or more plant-derived feed ingredients, substantially unsupplemented  
 CC with small-molecule antibiotics and as an additive a seed composition  
 CC containing a flour, extract or malt obtained from mature monocot seeds  
 CC and one or more heterologous seed-produced anti-microbial proteins in  
 CC substantially unpurified form. The invention is useful as a feed for  
 CC production animals such as poultry and hoofed farm animals. The present  
 CC sequence is human codon optimised lactoferrin. This sequence is used in  
 CC the invention

SO Sequence 690 AA;

Query Match 79.0%; Score 39.5; DB 5; Length 690;  
 Best Local Similarity 90.9%; Pred. No. 69;  
 Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RAVVM 10



Db 336 EVAARRARVW 346

RESULT 13  
AA58733  
ID AA58733 standard; protein; 692 AA.

XX AC AA58733;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin.

XX KW Lactoferrin; human; transgenic plant.

XX OS Homo sapiens.

XX PN WO200004146-A1.

XX PD 27-JAN-2000.

XX PF 19-JUL-1999; 99WO-IT000226.

XX PR 17-JUL-1998; 98IT-RM000478.

XX PA (PLAN-) PLANTECHNO SRL.

XX PI Fogher C;

XX DR WPI; 2000-161129/14.

XX DR N-PSDB; AA258122.

PT Synthetic polynucleotide encoding human lactoferrin, used for production of functional foods, vegetal milks and human lactoferrin.

PS Disclosure; Page 73-77; 93pp; English.

XX CC The present sequence is that of human lactoferrin. The invention relates to a synthetic gene (see AA258122) that encodes human lactoferrin but which has codon usage designed to maximise expression in plants.  
XX CC Transgenic plants that express human lactoferrin in a tissue-specific manner, especially in the seeds, can be used in processes for the production of functional vegetal milk, fruit juices, fruit and/or vegetable homogenized foods (claimed). The transgenic plants are selected from solanaceae, cereals, leguminosae, fruit bearing plants and horticultural plants, especially soybean, tobacco and rice

SQ Sequence 692 AA;

Query Match 79.0%; Score 39.5; DB 3; Length 692;  
Best Local Similarity 90.9%; Pred. No. 69;

Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAAR-RARVW 10

Db 338 EVAARRARVW 348

RESULT 14  
AAB97382

ID AAB97382 standard; protein; 692 AA.

XX AC AAB97382;

XX DT 17-AUG-2001 (first entry)

XX DE Human lactoferrin (hLF).

XX KW Human; lactoferrin; hLF; N-terminal; antimicrobial; heparin; inflammatory response; cytokine production reduction; neutrophil degranulation; myelopoiesis inhibition.

OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Region 1..27 /note= "Fragments of the N-terminal are specifically referred to in the claims"

XX FT Domain 2..5 /label= Cationic\_domain

XX FT Domain 28..31 /label= Cationic\_domain

XX PN WO200134641-A2.

XX PD 17-MAY-2001.

XX PF 10-NOV-2000; 2000WO-NL000821.

XX PR 11-NOV-1999; 99EP-00203775.

XX PR 11-NOV-1999; 99US-0164975P.

XX PA (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.

XX PI Van Berkel PHC, Nibbering PH, Nuijens JH;

XX DR WPI; 2001-335909/35.

XX PT New polypeptides comprising the N-terminal region of human lactoferrin protein exhibit higher antimicrobial activity than the full length protein and are useful to treat bacterial infections.

XX PS Claim 1; Page 55-57; 59pp; English.

XX CC This invention relates to fragments (between 6 and 26 amino acids) of the human lactoferrin hLF protein (represented by the present sequence). N-terminal hLF peptides have antimicrobial activity. The peptides of the invention are used to treat microbial infections, especially infections by gram positive or negative bacteria, particularly *Listeria*, *Staphylococcus*, *Klebsiella* or *Escherichia* species, especially *L. monocytogenes*, *S. aureus* and *E. coli*. Other uses include reducing inflammatory response by neutralising heparin or lipopolysaccharide or by reducing cytokine production and neutrophil degranulation, inhibiting entry of viruses such as cytomegalovirus, herpes simplex virus 1 or HIV into cells, inhibiting myelopoiesis and reducing production of GM-CSF (granulocyte/macrophage colony stimulating factor)

SQ Sequence 692 AA;

Query Match 79.0%; Score 39.5; DB 4; Length 692;  
Best Local Similarity 90.9%; Pred. No. 69;

Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAAR-RARVW 10

Db 338 EVAARRARVW 348

RESULT 15  
AAG77906

ID AAG77906 standard; protein; 692 AA.

XX AC AAG77906;

XX DT 18-JAN-2002 (first entry)

XX DE Human lactoferrin.

XX KW Human; lactoferrin; hLF; infectious disease; inflammatory disease; excess of heparin; gastroenteritis; inflammatory bowel disease; sepsis; anaemia; myelopoiesis; reperfusion injury; cytokine release; proteoglycan.

XX OS Homo sapiens.

```

FH Key Location/Qualifiers
FT Domain 2..5 /label= Cationic_domain
FT Domain 28..31 /label= Cationic_domain
FT Domain /label= Cationic_domain
XX
XX WO200172322-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-NL000253.
XX
XX 27-MAR-2000; 2000EP-00201110.
XX
XX 27-MAR-2000; 2000US-0193352P.
XX
XX (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.
XX
XX Van Bree JBM, Nuijens JH;
XX
XX WPI; 2001-648424/74.
XX
XX
XX Use of lactoferrin for treatment of infectious diseases, inflammatory
XX diseases and excess of heparin.
XX
XX Claim 26; Page 47-49; 49p; English.
XX
XX The sequence represents the human lactoferrin (hLF) protein. The methods
XX of the invention are useful for the treatment or prophylaxis of
XX infectious diseases, inflammatory diseases and excess of heparin e.g.
XX gastroenteritis, inflammatory bowel disease, sepsis, anaemia,
XX myeloperoxidases, reducing reperfusion injury, cytokine release and
XX proteoglycan-mediated entry of virus into cells. The advantage of the
XX method is that the patient is substantially free of side effect responses
XX to administration of lactoferrin. Therefore large doses of lactoferrin
XX can be administered
XX
XX SQ Sequence 692 AA;
XX
XX Query Match 79.0%; Score 39.5; DB 4; Length 692;
XX Best Local Similarity 90.9%; Pred. No. 69;
XX Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 EVAA-RAAYW 10
   |||||
   338 EVAARRARVW 348

```

Search completed: September 1, 2004, 00:09:43  
 Job time : 28.0667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2004, 00:04:10 : Search time 6.46667 Seconds  
(without alignments)  
79.834 Million cell updates/sec

Title: US-09-508-095-8  
Perfect score: 50  
Sequence: 1 EVAAARAVVW 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCITUS-COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39.5	79.0	694	US-08-724-586-2	Sequence 2, Appli
2	39.5	79.0	694	US-09-421-632-2	Sequence 2, Appli
3	39.5	79.0	694	US-09-932-190-2	Sequence 2, Appli
4	39.5	79.0	705	US-08-655-640-2	Sequence 2, Appli
5	39.5	79.0	708	US-08-655-640-4	Sequence 4, Appli
6	39.5	79.0	709	US-08-154-019-2	Sequence 2, Appli
7	39.5	79.0	709	US-08-461-333-2	Sequence 2, Appli
8	39.5	79.0	709	US-08-464-167-2	Sequence 2, Appli
9	39.5	79.0	709	US-09-158-313-2	Sequence 2, Appli
10	39.5	79.0	709	US-08-476-798-2	Sequence 2, Appli
11	39.5	79.0	711	US-08-145-681-2	Sequence 2, Appli
12	39.5	79.0	711	US-08-250-308-2	Sequence 2, Appli
13	39.5	79.0	711	US-08-154-019-4	Sequence 4, Appli
14	39.5	79.0	711	US-08-461-333-4	Sequence 4, Appli
15	39.5	79.0	711	US-08-453-703-2	Sequence 2, Appli
16	39.5	79.0	711	US-08-456-106-2	Sequence 2, Appli
17	39.5	79.0	711	US-08-464-167-4	Sequence 4, Appli
18	39.5	79.0	711	US-09-158-313-4	Sequence 4, Appli
19	39.5	79.0	711	US-08-456-108-2	Sequence 2, Appli
20	39.5	79.0	711	US-08-476-798-4	Sequence 4, Appli
21	39.5	79.0	711	US-09-265-577-2	Sequence 2, Appli
22	39.5	79.0	711	US-09-633-739-2	Sequence 2, Appli
23	39.5	79.0	711	PCT-US93-03614-2	Sequence 2, Appli
24	68.0	973	4	US-09-252-991A-21386	Sequence 21386, A
25	66.0	426	3	US-09-320-878-8	Sequence 6, Appli
26	66.0	426	3	US-09-105-537-20	Sequence 20, Appli
27	66.0	426	4	US-09-141-908-8	Sequence 8, Appli

28	33	66.0	426	4	US-09-657-440-8	Sequence 8, Appli
29	33	66.0	769	4	US-09-252-991A-19611	Sequence 19611, A
30	33	66.0	852	4	US-08-811-682-3	Sequence 3, Appli
31	33	66.0	3546	4	US-09-679-279-13	Sequence 13, Appli
32	33	66.0	3782	3	US-09-105-537-4	Sequence 4, Appli
33	32	64.0	124	3	US-09-732-210-357	Sequence 357, App
34	32	64.0	151	4	US-09-252-991A-24960	Sequence 24960, A
35	32	64.0	245	4	US-09-488-039A-12567	Sequence 12567, A
36	32	64.0	646	4	US-09-488-039A-12750	Sequence 12750, A
37	32	64.0	667	4	US-09-196-270-2	Sequence 2, Appli
38	32	64.0	804	3	US-08-909-954-2	Sequence 2, Appli
39	31.5	63.0	703	1	US-08-145-681-6	Sequence 6, Appli
40	31.5	63.0	703	1	US-08-453-703-6	Sequence 6, Appli
41	31.5	63.0	703	2	US-08-456-106-6	Sequence 6, Appli
42	31.5	63.0	703	3	US-08-456-108-6	Sequence 6, Appli
43	31.5	63.0	703	3	US-09-265-577-6	Sequence 6, Appli
44	31.5	63.0	703	4	US-09-633-739-6	Sequence 6, Appli
45	31	62.0	97	4	US-09-252-991A-32559	Sequence 32559, A

## ALIGNMENTS

RESULT 1  
US-08-724-586-2  
Sequence 2, Application US/08724586  
Patent No. 6064469  
GENERAL INFORMATION:  
APPLICANT: Krutzel, Marian L.  
APPLICANT: Kurecki, Tomas  
APPLICANT: Golinski, Paul D.  
APPLICANT: Doyle, Darrell J.  
TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Jacobson, Price, Holman & Stern  
STREET: 400 Seventh St. N.W.  
CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,586  
FILING DATE: 30-SEPT-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/238,445  
FILING DATE: 05-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: 10505/P58185C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-724-586-2  
Query Match 79.0%; Score 39.5; DB 3; Length 694;  
Best Local Similarity 90.9%; Pred. No. 21;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 EVAA-RARVW 10

DB 340 EVAAARRRVVW 350

## RESULT 2

US-09-421-632-2  
Sequence 2, Application US/09421632

Patent No. 6277817

GENERAL INFORMATION:

APPLICANT: Krutzel, Marian L.

APPLICANT: Kurecki, Tomasz

APPLICANT: Gollnick, Paul D.

APPLICANT: Doyle, Darrell J.

TITLE OF INVENTION: Cloning, Expression, and Uses of Human

TITLE OF INVENTION: Lactoferrin

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jacobson, Price, Holman & Stern

STREET: 400 Seventh St. N.W.

CITY: Washington D.C.

COUNTRY: U.S.A.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/421,632

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/724,586

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Player, William E.

REGISTRATION NUMBER: 31,409

REFERENCE/DOCKET NUMBER: 10505/P58185C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-6666

TELEFAX: (202) 393-5350

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 694 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-421-632-2

Query Match 79.0%; Score 39.5; DB 3; Length 694;  
Best Local Similarity 90.9%; Pred. No. 21;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 EVAA-RRRVVW 10  
DB 340 EVAAARRRVVW 350

## RESULT 3

US-09-932-190-2  
Sequence 2, Application US/09932190

Patent No. 6455687

GENERAL INFORMATION:

APPLICANT: Krutzel, Marian L.

APPLICANT: Kurecki, Tomasz

APPLICANT: Gollnick, Paul D.

APPLICANT: Doyle, Darrell J.

TITLE OF INVENTION: Cloning, Expression, and Uses of Human

TITLE OF INVENTION: Lactoferrin

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jacobson, Price, Holman & Stern

STREET: 400 Seventh St. N.W.

CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/932,190

FILING DATE: 17-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/724,586

FILING DATE: 30-SEPT-1996

APPLICATION NUMBER: US 08/238,445

FILING DATE: 05-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: Player, William E.

REGISTRATION NUMBER: 31,409

REFERENCE/DOCKET NUMBER: 10505/P58185C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-6666

TELEFAX: (202) 393-5350

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 694 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-932-190-2

Query Match 79.0%; Score 39.5; DB 4; Length 694;  
Best Local Similarity 90.9%; Pred. No. 21;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 EVAA-RRRVVW 10  
DB 340 EVAAARRRVVW 350

## RESULT 4

US-08-655-640-2  
Sequence 2, Application US/08655640

Patent No. 5948613

GENERAL INFORMATION:

APPLICANT: Teng, Christina

APPLICANT: Panella, Timothy J.

TITLE OF INVENTION: HUMAN LACTOFERRIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DARBY & CUSHMAN

STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/655,640

FILING DATE: 30-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/992,538

FILING DATE: December 17, 1992

ATTORNEY/AGENT INFORMATION:

NAME: SCOTT, WATSON T.

REGISTRATION NUMBER: 26,581

REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 705 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-655-640-2

Query Match 79.0%; Score 39.5; DB 2; Length 705;

Best Local Similarity 90.9%; Pred. No. 22;

Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10

DB 355 EVAARRARYVM 365

RESULT 5

US-08-655-640-4

Sequence 4, Application US/08655640

Patent No. 5948613

GENERAL INFORMATION:

APPLICANT: Teng, Christina

APPLICANT: Panella, Timothy J.

TITLE OF INVENTION: HUMAN LACTOFERRIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CUSHMAN, DARBY & CUSHMAN

STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/655,640

FILING DATE: 30-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/992,538

FILING DATE: December 17, 1992

ATTORNEY/AGENT INFORMATION:

NAME: SCOTT, WATSON T.

REGISTRATION NUMBER: 26,581

REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 861-3000

TELEFAX: (202) 822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 708 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-655-640-4

Query Match 79.0%; Score 39.5; DB 2; Length 708;

Best Local Similarity 90.9%; Pred. No. 22;

Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10

DB 357 EVAARRARYVM 367

RESULT 6

US-08-154-019-2

Sequence 2, Application US/08154019

Patent No. 5633076

GENERAL INFORMATION:

APPLICANT: Deboer, Herman A.

APPLICANT: Strijker, Rein

APPLICANT: Heyneker, Herbert L.

APPLICANT: Platenburg, Gerald

APPLICANT: Lee, Sang He

APPLICANT: Pieper, Frank

APPLICANT: Krimpenfort, Paul J. A.

TITLE OF INVENTION: Production of Recombinant Polypeptides

by Bovine Species and Transgenic Methods

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Townsend and Crew

STREET: One Market Plaza, Stewart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/154,019

FILING DATE: 16-NOV-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/461,333

FILING DATE: 05-JUN-1995

APPLICATION NUMBER: US 08/077,788

FILING DATE: 15-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/695,956

FILING DATE: 15-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/619,131

FILING DATE: 27-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/444,745

FILING DATE: 01-DEC-1989

ATTORNEY/AGENT INFORMATION:

NAME: Liebescheutz, Joe O.

REGISTRATION NUMBER: 37,505

REFERENCE/DOCKET NUMBER: 16994-003123

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-543-5043

TELEFAX: 415-543-5043

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 709 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-154-019-2

Query Match 79.0%; Score 39.5; DB 1; Length 709;

Best Local Similarity 90.9%; Pred. No. 22;

Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10

DB 355 EVAARRARYVM 365

RESULT 7

US-08-461-333-2

Sequence 2, Application US/08461333  
Patent No. 5741957  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,333  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschultz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-461-333-2

Query Match 79.0%; Score 39.5; DB 1; Length 709;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10  
Db 355 EVAARRARYVM 365

RESULT 8  
US-08-464-167-2  
Sequence 2, Application US/08464167  
Patent No. 6013857  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.

APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,167  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschultz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003124  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-167-2

Query Match 79.0%; Score 39.5; DB 3; Length 709;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10  
Db 355 EVAARRARYVM 365

RESULT 9  
US-09-158-313-2  
Sequence 2, Application US/09158313  
Patent No. 6066725  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods

NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM PC floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/158,313  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,798  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebescheutz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-158-313-2

Query Match 79.0%; Score 39.5; DB 3; Length 709;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10  
DB 355 EVAARRARYVM 365

RESULT 10  
US-08-476-798-2  
Sequence 2, Application US/08476798  
Patent No. 6140552  
GENERAL INFORMATION:  
APPLICANT: DeBoer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California

COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,798  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebescheutz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-798-2

Query Match 79.0%; Score 39.5; DB 3; Length 709;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10  
DB 355 EVAARRARYVM 365

RESULT 11  
US-08-145-681-2  
Sequence 2, Application US/08145681  
Patent No. 5571691  
GENERAL INFORMATION:  
APPLICANT: Conneely, Orla M.  
APPLICANT: Heaton, Denis R.  
APPLICANT: O'Malley, Bert W.  
APPLICANT: May, Gregory S.  
TITLE OF INVENTION: Production of Recombinant Lactoferrin  
TITLE OF INVENTION: and Lactoferrin Polypeptides Using cDNA Sequences in  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Baker & Botts, L.L.P.  
STREET: 910 Louisiana St  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77002  
COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM PC floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/145,681  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcgregor, Martin L.  
REGISTRATION NUMBER: 29,329  
REFERENCE/DOCKET NUMBER: 19928-0125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713/229/1874  
TELEFAX: 713/229/1522  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: H. sapiens  
US-08-145-681-2

Query Match 79.0%; Score 39.5; DB 1; Length 711;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10  
DB 357 EVAAARRARYVM 367

RESULT 12  
US-08-250-308-2  
Sequence 2, Application US/08250308  
Patent No. 5571896  
Patent No. 5571896 5571896  
GENERAL INFORMATION:  
APPLICANT: Connealy, Orla M.  
APPLICANT: Headon, Denis R.  
APPLICANT: O'Malley, Bert W.  
TITLE OF INVENTION: Production of Recombinant Human  
TITLE OF INVENTION: Lactoferrin  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fulbright & Jaworski Patent Department  
STREET: 1301 McKinney St.  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77010-3095  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/250,308  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/873,304  
FILING DATE: 24-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler Ph.D., Benjamin A.  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5456  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713/6515587  
TELEFAX: 713/6515246  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-250-308-2

Query Match 79.0%; Score 39.5; DB 1; Length 711;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARYVM 10  
DB 357 EVAAARRARYVM 367

RESULT 13  
US-08-154-019-4  
Sequence 4, Application US/08154019  
Patent No. 5633076  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Platenburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/154,019  
FILING DATE: 16-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/461,333  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschuetz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-154-019-4



Query Match 79.0%; Score 39.5; DB 1; Length 711;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10  
DB 357 EVAARRARVW 367

RESULT 14  
US-08-461-333-4  
Sequence 4, Application US/08461333  
Patent No. 5741957  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
APPLICANT: Strijker, Rein  
APPLICANT: Heyneker, Herbert L.  
APPLICANT: Plateburg, Gerald  
APPLICANT: Lee, Sang He  
APPLICANT: Pieper, Frank  
APPLICANT: Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
TITLE OF INVENTION: by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,333  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschultz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-461-333-4

Query Match 79.0%; Score 39.5; DB 1; Length 711;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10  
DB 357 EVAARRARVW 367

DB 357 EVAARRARVW 367

RESULT 15  
US-08-453-703-2  
Sequence 2, Application US/08453703  
Patent No. 576939  
GENERAL INFORMATION:  
APPLICANT: Connelly, Orla M.  
APPLICANT: Headon, Denis R.  
APPLICANT: O'Malley, Bert W.  
APPLICANT: May, Gregory S.  
TITLE OF INVENTION: Production of Recombinant Lactoferrin  
TITLE OF INVENTION: and Lactoferrin Polypeptides Using cDNA Sequences in  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,703  
FILING DATE: Concurrently herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/145,681  
FILING DATE: October 28, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Albert P. Halluin  
REGISTRATION NUMBER: 25,227  
REFERENCE/DOCKET NUMBER: 8206-024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-3660  
TELEFAX: 415-854-3694  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: H. sapiens  
US-08-453-703-2

Query Match 79.0%; Score 39.5; DB 1; Length 711;  
Best Local Similarity 90.9%; Pred. No. 22;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10  
DB 357 EVAARRARVW 367

Search completed: September 1, 2004, 00:18:37  
Job time : 7.46667 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2004, 00:15:50 ; Search time 22.8 Seconds

(without alignments)  
137.988 Million cell updates/sec

Title: US-09-508-095-8  
Perfect score: 50  
Sequence: 1 EVARRARVW 10

Scoring table: BLASTSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:\*  
1: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/2/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/2/pubppa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/prodata/2/pubppa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/2/pubppa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/prodata/2/pubppa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/prodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39.5	79.0	359	US-10-169-297-49	Sequence 49, Appli
2	39.5	79.0	680	US-10-076-816-4	Sequence 4, Appli
3	39.5	79.0	690	US-10-077-381-4	Sequence 4, Appli
4	39.5	79.0	690	US-10-639-835-4	Sequence 4, Appli
5	39.5	79.0	694	US-10-023-096-2	Sequence 2, Appli
6	39.5	79.0	709	US-10-170-221-2	Sequence 2, Appli
7	39.5	79.0	711	US-10-169-297-9	Sequence 2, Appli
8	39.5	79.0	711	US-10-170-221-4	Sequence 4, Appli
9	39.5	79.0	711	US-10-341-434-202	Sequence 202, App
10	39.5	79.0	711	US-10-341-434-218	Sequence 218, App
11	39.5	79.0	711	US-10-440-464-69	Sequence 69, Appli
12	39.5	79.0	711	US-10-440-765A-2000	Sequence 2000, Ap
13	38	76.0	110	US-10-437-963-169771	Sequence 169771,
14	35	70.0	153	US-10-437-963-126957	Sequence 126957,
15	35	70.0	254	US-10-156-761-10952	Sequence 10952, A

## ALIGNMENTS

RESULT 1  
US-10-169-297-49  
; Sequence 49, Application US/10169297  
; Publication No. US20030171276A1  
; GENERAL INFORMATION:  
; APPLICANT: Tohdo, Naoki  
; APPLICANT: Murata, Masashi  
; APPLICANT: Enjoji, Takashi  
; TITLE OF INVENTION: Preventives and Remedies for Chronic  
; TITLE OF INVENTION: Hepatitis  
; FILE REFERENCE: 3435.1000-000  
; CURRENT APPLICATION NUMBER: US/10/169,297  
; PRIOR FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: PCT/JP00/09393  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: JP 11/374087  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 49  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-169-297-49  
Query Match 79.0%; Score 39.5; DB 14; Length 359;  
Best Local Similarity 90.9%; Pred. No. 36;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
DB 111 EVARRARVW 121  
RESULT 2  
US-10-076-816-4  
; Sequence 4, Application US/10076816  
; Publication No. US20030056244A1

```

; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,188
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-816-4
```

```

Query Match          79.0%; Score 39.5; DB 14; Length 690;
Best Local Similarity 90.9%; Pred. No. 67;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY      1 EVAA-RARVW 10
DB      336 EVAARRARVW 346
```

```

RESULT 3
US-10-077-381-4
; Sequence 4, Application US/10077381
; Publication No. US20030074700A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Expression of Human Milk Proteins in
; FILE REFERENCE: 50665-8022.US00
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,139
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-077-381-4
```

```

Query Match          79.0%; Score 39.5; DB 14; Length 690;
Best Local Similarity 90.9%; Pred. No. 67;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY      1 EVAA-RARVW 10
DB      336 EVAARRARVW 346
```

```

RESULT 4
US-10-639-835-4
```

```

; Sequence 4, Application US/10639835
; Publication No. US20040111766A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Expression of Human Milk Proteins in
; FILE REFERENCE: 50665-8022.US01
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: US 10/077,381
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,199
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-639-835-4
```

```

Query Match          79.0%; Score 39.5; DB 16; Length 690;
Best Local Similarity 90.9%; Pred. No. 67;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY      1 EVAA-RARVW 10
DB      336 EVAARRARVW 346
```

```

RESULT 5
US-10-023-096-2
; Sequence 2, Application US/10023096
; Publication No. US20020160941A1
; GENERAL INFORMATION:
; APPLICANT: Krutzel, Marian L.
; APPLICANT: Kurecki, Tomasz
; APPLICANT: Golnick, Paul D.
; APPLICANT: Doyle, Darrell J.
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human
; TITLE OF INVENTION: Lactoferrin
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern
; STREET: 400 Seventh St. N.W.
; CITY: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/023,096
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,586
; FILING DATE: 30-SEPT-1996
; APPLICATION NUMBER: US 08/238,445
; FILING DATE: 05-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409
```

REFERENCE/DOCKET NUMBER: 10505/PS8185C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 79.0%; Score 39.5; DB 13; Length 694;  
Best Local Similarity 90.9%; Pred. No. 68;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10  
Db 340 EVAA-RARVW 350

RESULT 6  
US-10-170-221-2  
Sequence 2, Application US/10170221  
Publication No. US20030192068A1  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
Strijker, Rein  
Heyneker, Herbert L.  
Platenburg, Gerald  
Lee, Sang He  
Pieper, Frank  
Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/170,221  
FILING DATE: 11-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,798  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/077,788  
FILING DATE: 15-JUN-1993  
APPLICATION NUMBER: US 07/895,956  
FILING DATE: 15-JUN-1992  
APPLICATION NUMBER: US 07/619,131  
FILING DATE: 27-NOV-1990  
APPLICATION NUMBER: US 07/444,745  
FILING DATE: 01-DEC-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Liedescheutz, Joe O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 16994-003125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-170-221-2

Query Match 79.0%; Score 39.5; DB 14; Length 709;  
Best Local Similarity 90.9%; Pred. No. 69;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10  
Db 355 EVAA-RARVW 365

RESULT 7  
US-10-169-297-9  
Sequence 9, Application US/10169297  
Publication No. US20030171276A1  
GENERAL INFORMATION:  
APPLICANT: Tomoda, Naoki  
Murai, Masashi  
Enjoji, Takashi  
TITLE OF INVENTION: Preventives and Remedies for Chronic  
Hepatitis  
FILE REFERENCE: 3435,1000-000  
CURRENT APPLICATION NUMBER: US/10/169,297  
CURRENT FILING DATE: 2002-10-31  
PRIOR APPLICATION NUMBER: PCT/JP00/09393  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: JP 11/374087  
PRIOR FILING DATE: 1999-12-28  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 711  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-169-297-9

Query Match 79.0%; Score 39.5; DB 14; Length 711;  
Best Local Similarity 90.9%; Pred. No. 69;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RARVW 10  
Db 357 EVAA-RARVW 367

RESULT 8  
US-10-170-221-4  
Sequence 4, Application US/10170221  
Publication No. US20030192068A1  
GENERAL INFORMATION:  
APPLICANT: Deboer, Herman A.  
Strijker, Rein  
Heyneker, Herbert L.  
Platenburg, Gerald  
Lee, Sang He  
Pieper, Frank  
Krimpenfort, Paul J.A.  
TITLE OF INVENTION: Production of Recombinant Polypeptides  
by Bovine Species and Transgenic Methods  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/170,221
FILING DATE: 11-Jun-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/476,798
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/077,788
FILING DATE: 15-JUN-1992
APPLICATION NUMBER: US 07/895,956
FILING DATE: 15-JUN-1992
APPLICATION NUMBER: US 07/619,131
FILING DATE: 27-NOV-1990
APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-170-221-4

Query Match          79.0%; Score 39.5; DB 14; Length 711;
Best Local Similarity 90.9%; Pred. No. 69;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RAVVW 10
DB 357 EVAARRARVVW 367

RESULT 9
US-10-341-434-202
Sequence 202, Application US/10341434
Publication No. US20030215835A1
GENERAL INFORMATION:
APPLICANT: Oxigene Technologies
TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
FILE REFERENCE: 9U 204 205 R1
CURRENT APPLICATION NUMBER: US/10/341,434
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: US 60/348,164
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: US 60/348,119
PRIOR FILING DATE: 2002-01-15
NUMBER OF SEQ ID NOS: 238
SOFTWARE: Patentin version 3.1
SEQ ID NO 202
LENGTH: 711
TYPE: PRT
ORGANISM: Homo sapiens
US-10-341-434-202

Query Match          79.0%; Score 39.5; DB 15; Length 711;
Best Local Similarity 90.9%; Pred. No. 69;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RAVVW 10
DB 357 EVAARRARVVW 367
```

```
RESULT 10
US-10-341-434-218
Sequence 218, Application US/10341434
Publication No. US20030215835A1
GENERAL INFORMATION:
APPLICANT: Oxigene Technologies
TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
FILE REFERENCE: 9U 204 205 R1
CURRENT APPLICATION NUMBER: US/10/341,434
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: US 60/348,164
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: US 60/348,119
PRIOR FILING DATE: 2002-01-15
NUMBER OF SEQ ID NOS: 238
SOFTWARE: Patentin version 3.1
SEQ ID NO 218
LENGTH: 711
TYPE: PRT
ORGANISM: Homo sapiens
US-10-341-434-218

Query Match          79.0%; Score 39.5; DB 15; Length 711;
Best Local Similarity 90.9%; Pred. No. 69;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RAVVW 10
DB 357 EVAARRARVVW 367

RESULT 11
US-10-440-464-69
Sequence 69, Application US/10440464
Publication No. US20040018528A1
GENERAL INFORMATION:
APPLICANT: DEPRIMO, SAMUEL
APPLICANT: O'FARRELL, ANNE-MARIE
APPLICANT: MORIMOTO, ALYSSA
APPLICANT: SMOICH, BEVERLY
APPLICANT: MANNING, WILLIAM
APPLICANT: WALTER, SARAH
APPLICANT: CHERRINGTON, JULIE
TITLE OF INVENTION: NOVEL BIOMARKERS OF TYROSINE KINASE INHIBITOR EXPOSURE
TITLE OF INVENTION: AND ACTIVITY IN MAMMALS
FILE REFERENCE: 038602/1592
CURRENT APPLICATION NUMBER: US/10/440,464
CURRENT FILING DATE: 2003-05-19
PRIOR APPLICATION NUMBER: 60/380,872
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: 60/448,922
PRIOR FILING DATE: 2003-02-24
PRIOR APPLICATION NUMBER: 60/448,874
PRIOR FILING DATE: 2003-02-24
NUMBER OF SEQ ID NOS: 185
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 69
LENGTH: 711
TYPE: PRT
ORGANISM: Homo sapiens
US-10-440-464-69

Query Match          79.0%; Score 39.5; DB 15; Length 711;
Best Local Similarity 90.9%; Pred. No. 69;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EVAA-RAVVW 10
DB 357 EVAARRARVVW 367
```

```
RESULT 12
US-10-408-765A-2000
; Sequence 2000, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Pany, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088,465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2000
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2000

Query Match      79.0%; Score 39.5; DB 16; Length 711;
Best Local Similarity 90.9%; Pred. No. 69;
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy      1 EVAARARVW 10
Db      357 EVAARARVW 367

RESULT 13
US-10-437-963-169771
; Sequence 169771, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 169771
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68159C.1.pcp
US-10-437-963-169771

Query Match      76.0%; Score 38; DB 16; Length 110;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      3 AARARVW 10
Db      48 AARARVW 55

RESULT 14
US-10-437-963-126957
; Sequence 126957, Application US/10437963

; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126957
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(153)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: unsure at all Xaa locations
US-10-437-963-126957

Query Match      70.0%; Score 35; DB 16; Length 153;
Best Local Similarity 70.0%; Pred. No. 11e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 EVAARARVW 10
Db      14 EVARRSRDVM 23

RESULT 15
US-10-156-761-10952
; Sequence 10952, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHISA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10952
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10952

Query Match      70.0%; Score 35; DB 14; Length 254;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      2 VAAARVW 10
Db      32 IARRARVW 40
```

Wed Sep 1 08:24:15 2004

us-09-508-095-8.aug31.rapb

Search completed: September 1, 2004, 00:49:16  
Job time : 24.8 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2004, 00:03:45 ; Search time 5.8 seconds

(without alignments)  
165.847 Million cell updates/sec

Title: US-09-508-095-8

Sequence: 1 EVVARARVW 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

PIR 78: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39.5	79.0	711	1	TEFHUL
2	39	78.0	175	2	A43815
3	38	76.0	912	2	B75127
4	36	72.0	1918	2	S43719
5	36	72.0	1920	2	S43721
6	36	72.0	1926	2	S01169
7	36	72.0	1937	2	S01169
8	35	70.0	108	2	F72507
9	35	70.0	210	2	D87507
10	35	70.0	380	2	A12736
11	35	70.0	380	2	G97517
12	35	70.0	392	2	S72753
13	35	70.0	689	2	A95418
14	35	70.0	1337	2	T13948
15	34	68.0	117	2	JN0731
16	34	68.0	146	2	T35484
17	34	68.0	158	2	T46199
18	34	68.0	213	2	A87259
19	34	68.0	387	2	B70325
20	34	68.0	390	2	S76171
21	34	68.0	706	2	D82160
22	34	68.0	707	1	A28438
23	34	68.0	855	2	F45557
24	34	68.0	930	2	D37271
25	34	68.0	1458	2	T17402
26	33	66.0	285	1	IDECRP
27	33	66.0	285	1	I64780
28	33	66.0	302	2	AB0780
29	33	66.0	331	2	AF1349

#### ALIGNMENTS

##### RESULT 1

TEFHUL

Lactotransferrin precursor [validated] - human

N/Alternate names: lactoferrin

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C/Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S7

R/Cho: Y.

submitted to the EMBL Data Library, March 1994

A/Reference number: G06820

A/Accession: G01394

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-711 <CHO>

A/Cross-references: EMBL:U07643; NID:9467236; PIDN:AAB60324.1; PID:9467237

R/Rev: M.W.; Moloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A/Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A/Reference number: S11228; MUID:90384839; PMID:2402455

A/Accession: S11228

A/Molecule type: mRNA

A/Residues: 1-148, 'T', 150-422, 'C', 424-711 <REV>

A/Cross-references: EMBL:X53961; NID:934415; PIDN:CA37914.1; PID:934416

R/Teng, C.T.; Liu, Y.; Yang, N.; Palmer, D.; Pannella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A/Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A/Reference number: A45401; MUID:93125571; PMID:1480183

A/Accession: A45401

A/Molecule type: DNA

A/Residues: 1-15 <DNA>

A/Cross-references: GB:S52659; NID:9263311; PIDN:AAB24877.1; PID:9263312

A/Experimental source: placenta

A/Note: sequence extracted from NCBI backbone (NCBI:122202)

R/Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A/Title: Nucleotide sequence of human lactoferrin cDNA.

A/Reference number: S10324; MUID:90326549; PMID:2374734

A/Accession: S10324

A/Molecule type: mRNA

A/Residues: 3-711 <POM>

A/Cross-references: EMBL:X52941; NID:934411; PIDN:CA37116.1; PID:934412

R/Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A/Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A/Reference number: S15853; MUID:91264786; PMID:2049066

A/Accession: S15853

A/Status: nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

tryptophanyl-tRNA  
hypothetical prote  
conserved hypothet  
pho4 family protei  
hypothetical prote  
hypothetical prote  
hypothetical prote  
probable primase Y  
hypothetical prote  
hypothetical prote  
hypothetical prote  
adaptor protein/a  
beta3 protein - fr  
hypothetical prote  
ribosomal protein  
acetyltransferase,

R:Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A>Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A/Reference number: S07160; MUID:86001031; PMID:3477300  
 A/Accession: S07160  
 A/Molecule type: mRNA  
 A/Residues: 436-487, 'A', 489-711 <RAD>  
 A/Cross-references: EMBL:M8642; NID:9186815; PIND:AAA6665.1; PID:9386855  
 R:Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A>Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A/Reference number: A61169; MUID:91235214; PMID:1674448  
 A/Accession: A61169  
 A>Status: not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 3-701, 'SMKPVN' <PAN>  
 A/Experimental source: normal breast tissue  
 R:Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoengen, F.; Legrand, D.; Spik, G.; Eut, J.  
 Biochem. 143, 659-666, 1984  
 A>Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A/Reference number: A31000; MUID:85076667; PMID:6510420  
 A/Accession: A31000  
 A/Molecule type: protein  
 A/Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A/Note: this is the final paper in a series  
 R:Houen, G.; Hoegdall, E.V.; Barkholt, V.; Nørskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A>Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity  
 A/Reference number: S74119; MUID:97054624; PMID:8898921  
 A/Accession: S74119  
 A/Molecule type: protein  
 A/Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A/Experimental source: neutrophil granulocytes  
 A/Genetics:  
 A/Gene: GDB:LTF  
 A/Cross-references: GDB:119368; OMIM:150210  
 A/Map position: 3q21-3q23  
 A/Superfamily: transferrin; transferrin repeat homology  
 C/Keywords: duplication; glycoprotein; iron binding; milk  
 F:1-19/Domain: signal sequence #status predicted <SIC>  
 F:20-711/Product: lactotransferrin #status experimental <MAR>  
 F:21-356/Domain: transferrin repeat homology <TRH1>  
 F:360-699/Domain: transferrin repeat homology <TRH2>  
 F:29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-609/Disulfide bonds: #status experimental  
 F:157-498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:358-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #strat

Query Match 79.0%; Score 39.5; DB 1; Length 711;  
 Best Local Similarity 90.9%; Pred. No. 4.9; 0; Indels 1; Gaps 1;  
 Matches 10; Conservative 0; Mismatches

QY 1 EVAA-RRRVVW 10  
 DB 357 EVARRRRVW 367

RESULT 2  
 S43815  
 C/Species: phage SP1  
 C/Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 11-May-2000  
 C/Accession: S43815; T42345; S41189  
 R:Pedre, X.; Weise, F.; Chal, S.; Lueder, G.; Alonso, J.C.  
 J. Mol. Biol. 236, 1324-1340, 1994  
 A>Title: Analysis of cis and trans acting elements required for the initiation of DNA re  
 A/Reference number: S43798; MUID:4172631; PMID:8126723  
 A/Accession: S43815  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-175 <PBD>  
 A/Cross-references: EMBL:X67865; NID:9472886; PIND:CAA48064.1; PID:9439646  
 R:Alonso, J.C.; Luder, G.; Stiege, A.C.; Chal, S.; Weise, F.; Trautner, T.A.  
 Gene 204, 201-212, 1997

A>Title: The complete nucleotide sequence and functional organization of *Bacillus subtilis*  
 A/Reference number: Z22137; MUID:98094274; PMID:94334185  
 A/Accession: T42345  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-175 <ALO>  
 A/Cross-references: EMBL:X97918; PIND:CAA66499.1  
 C/Superfamily: phage SP1 hypothetical protein 41

Query Match 78.0%; Score 39; DB 2; Length 175;  
 Best Local Similarity 60.0%; Pred. No. 1.5;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVARRRVVW 10  
 DB 106 EISAKCRVW 115

RESULT 3  
 B75127  
 C/Species: *Pyrococcus abyssi*  
 C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 17-Nov-2000  
 C/Accession: B75127  
 R:Anonymous, Genoscope  
 A/Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru  
 A/Reference number: A75001  
 A/Accession: B75127  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-912 <KAW>  
 A/Cross-references: GB:AJ248285; GB:AL096836; NID:95458067; PIND:CAA9731.1; PID:e151562  
 A/Experimental source: strain Orsay  
 A/Genetics:  
 A/Gene: PAB1817

Query Match 76.0%; Score 38; DB 2; Length 912;  
 Best Local Similarity 70.0%; Pred. No. 13;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVARRRVVW 10  
 DB 230 EVARRRRTW 239

RESULT 4  
 S43719  
 C/Species: *Pyrococcus abyssi* / glycosylceramidase (EC 3.2.1.62) (clone BL20) - rabbit (fragmen  
 N/Alternate names: lactase / phloretin hydrolase  
 C/Species: *Oryctolagus cuniculus* (domestic rabbit)  
 C/Date: 13-Jan-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
 C/Accession: S43719  
 R:Villa, M.; Brunschweiler, D.; Gaechter, T.; Boll, W.; Semenza, G.; Mantel, N.  
 FEBS Lett. 336, 70-74, 1993  
 A>Title: Region-specific expression of multiple lactase-phloretin hydrolase genes in int  
 A/Reference number: S43719; MUID:94085594; PMID:8622219  
 A/Accession: S43719  
 A/Status: preliminary; translation not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-1918 <VIL>  
 A/Cross-references: EMBL:Z27166  
 C/Keywords: carbohydrate digestion; glycosidase; hydrolase; intestine

Query Match 72.0%; Score 36; DB 2; Length 1918;  
 Best Local Similarity 60.0%; Pred. No. 71;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVARRRVVW 10  
 DB 866 EVPSARRVW 875

```

RESULT 5
543721
lactase (EC 3.2.1.108) / glycosylceramidase (EC 3.2.1.62) (clone BL70) - rabbit (fragment)
N/Alternate names: lactase / phlorizin hydrolase
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 13-Jan-1995 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
J/Accession: S43720; S43721
R/Villa, M.; Brunschweiler, D.; Gaechter, T.; Boll, W.; Semenza, G.; Mantel, N.
FEBS Lett. 336, 70-74, 1993
A>Title: Region-specific expression of multiple lactase-phlorizin hydrolase genes in immature
A/Reference number: S43719; MUID:94085594; PMID:8262219
A/Accession: S43720
A>Status: preliminary; translation not shown
A/Molecule type: mRNA
A/Residues: 1-1920 <VIL>
A/Cross-references: EMBL:Z27167; NID:G415864; PIDN:CAA01691.1; PID:G415865
A/Accession: S43721
A>Status: preliminary; translation not shown
A/Molecule type: DNA
A/Residues: 'MELFWS', 1-11, 'Q', 13-207 <VI2>
A/Cross-references: EMBL:Z27168; NID:G415866; PIDN:CAA01692.1; PID:G415867
C/Keywords: carbohydrate digestion; glycosidase; hydrolase; intestine

Query Match      72.0%; Score 36; DB 2; Length 1920;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 EVARARVYV 10
      ||:::|||
Db      867 EVPSKAKVW 876

RESULT 6
501169
beta-glycosidase complex precursor - rabbit
N/Contains: glycosylceramidase (EC 3.2.1.62); lactase (EC 3.2.1.108)
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 05-Nov-1999
J/Accession: S01169
R/Mantel, N.; Villa, M.; Enzler, T.; Wacker, H.; Boll, W.; James, P.; Hunziker, W.; Semer
EMBO J. 7, 2705-2713, 1988
A>Title: Complete primary structure of human and rabbit lactase-phlorizin hydrolase: imple
A/Reference number: S01168; MUID:85030634; PMID:2460343
A/Accession: S01169
A/Molecule type: mRNA
A/Residues: 1-1926 <MAN>
A/Cross-references: EMBL:X07995; NID:G1616; PIDN:CAA30802.1; PID:G1617
A/Note: The authors translated the codon GCC for residue 1551 as Gly and CCA for residue
A/Note: part of this sequence, including the amino end of the mature protein, was confir
C/Keywords: carbohydrate digestion; glycosidase; hydrolase; intestine; transmembrane prote
F.1-19/Domain: signal sequence #status predicted <PRO>
F.20-866/Domain: propeptide #status predicted <PRO>
F.1883-1926/Product: beta-glycosidase complex #status experimental <MAT>
F.1883-1901/Domain: transmembrane #status predicted <TMN>

Query Match      72.0%; Score 36; DB 2; Length 1926;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 EVARARVYV 10
      ||:::|||
Db      873 EVPSKAKVW 882

RESULT 7
501168
beta-glycosidase complex precursor - human
N/Contains: glycosylceramidase (EC 3.2.1.62); lactase (EC 3.2.1.108)
C/Species: Homo sapiens (man)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
J/Accession: S01168
R/Mantel, N.; Villa, M.; Enzler, T.; Wacker, H.; Boll, W.; James, P.; Hunziker, W.; Semer
EMBO J. 7, 2705-2713, 1988

```

```

A>Title: Complete primary structure of human and rabbit lactase-phlorizin hydrolase: im
A:Reference number: S01168; MUID:85030634; PMID:2460343
A:Accession: S01168
A:Molecule type: mRNA
A:Residues: 1-1927 <MAN>
A:Cross-references: EMBL:X07994; NID:G34399; PIDN:CAA0801.1; PID:G34400
C:Keywords: carbohydrate digestion; glycosidase; hydrolase; intestine; transmembrane pr
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-866/Domain: propeptide #status predicted <PRO>
F:867-1277/Product: beta-glycosidase complex #status predicted <MAT>
F:1883-1901/Domain: transmembrane #status predicted <TM>

Query Match          72.0%; Score 36; DB 2; Length 1927;
Best Local Similarity 60.0%; Pred. No. 71;
Matches              6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db              875 EVPSKARVW 884

OY      1 EVAAARVW 10
      ||::|||
      ||::|||

RESULT 8
F72507
hypothetical protein APE2036 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72507
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KAM>
A:Cross-references: DDBJ:AP000063; NID:G5105654; PIDN:BA81046.1; PID:di044832; PID:G51
C:Genetics:
A:Experimental source: strain K1
A:Gene: APE2036
C:Superfamily: Aeropyrum pernix hypothetical protein APE2036

Query Match          70.0%; Score 35; DB 2; Length 108;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches              6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      3 AARARVW 10
      |||||
      |||||

Db              63 AERRARVW 70

RESULT 9
D87507
glutathione S-transferase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 02-Aug-2002
C:Accession: D87507
R:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heideberg, J
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87507
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <STO>
A:Cross-references: GB:AE005673; NID:gl3423566; PIDN:AAK24056.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2085
C:Superfamily: hypothetical protein b0838

Query Match          70.0%; Score 35; DB 2; Length 210;

```

Best Local Similarity 70.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVARARVW 10  
DB 82 EPRARVW 91

## RESULT 10

AI2736  
DNA processing chain A [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C/Species: Agrobacterium tumefaciens

C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002

C/Accession: AI2736

R/Mod: D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.

erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McCall

Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, B.W.

A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A/Reference number: AB2577; MUID:21608550; PMID:11743193

A/Accession: AI2736

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-380 <KUR>

A/Cross-references: GB:AE008668; PIDN:AA42311.1; PID:g17739713; GSPDB:GN00186

A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Map position: circular chromosome

A/Map position: circular chromosome

Query Match 70.0%; Score 35; DB 2; Length 380;

Best Local Similarity 60.0%; Pred. No. 22;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVARARVW 10  
DB 12 ELSARVW 21

RESULT 11

G97517

DNA processing chain A (AF088896) [imported] - Agrobacterium tumefaciens (strain C58, CE

C/Species: Agrobacterium tumefaciens

C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002

C/Accession: G97517

R/Godder, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A/Reference number: A97359; MUID:21608551; PMID:11743194

A/Accession: G97517

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-380 <KUR>

A/Cross-references: GB:AE007869; PIDN:AAK87096.1; PID:g15156358; GSPDB:GN00169

C/Genetics:

A/Map position: circular chromosome

A/Map position: circular chromosome

Query Match 70.0%; Score 35; DB 2; Length 380;

Best Local Similarity 60.0%; Pred. No. 22;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVARARVW 10  
DB 12 ELSARVW 21

RESULT 12

S72753

hypothetical protein B1496\_C1\_154 - Mycobacterium leprae

C/Species: Mycobacterium leprae

C/Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 23-Mar-2001

C/Accession: S72753; T11012

R/Smith, D.R.; Robison, K.

Submitted to the EMBL Data Library, November 1993

A/Description: Mycobacterium leprae cosmid B1496.

A/Reference number: S72695

A/Accession: S72753

A/Molecule type: DNA

A/Residues: 1-392 <SMI>

A/Cross-references: EMBL:U00013; NID:G466868; PIDN:AAA17119.1; PID:G466874

R/Parkhill, J.; Barrall, B.G.; Rajandream, M.A.

Submitted to the EMBL Data Library, September 1997

A/Accession: T11012

A/Status: preliminary; translated from GB/EMBL/DD81

A/Molecule type: DNA

A/Residues: 1-392 <PAR>

A/Cross-references: EMBL:Z99125; NID:G2398663; PIDN:CAB16170.1; PID:e343547; PID:G239870

C/Genetics:

A/Map position: circular chromosome

A/Map position: circular chromosome

Query Match 70.0%; Score 35; DB 2; Length 392;

Best Local Similarity 70.0%; Pred. No. 23;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVARARVW 10  
DB 164 EAAARVW 173

RESULT 13

A95418

probable UvrD2 DNA helicase (EC 3.6.1.-) [imported] - Sinorhizobium meliloti (strain 102)

C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C/Accession: A95418

R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalmann, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A/Reference number: A95262; MUID:21396509; PMID:11981432

A/Accession: A95418

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-689 <KUR>

A/Cross-references: GB:AE006469; PIDN:AAK65907.1; PID:g14524418; GSPDB:GN00165

A/Experimental source: strain 1021, megaplasmid pSymA

R/Gallbert, F.; Finan, T.M.; Long, S.R.; Punter, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalmann, S.; Keating, D.H.; Kiss, E.; Komp, C.; Leilaure,

hebaull, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; MUID:21368234; PMID:11974104

A/Contents: annotation

C/Genetics:

A/Map position: circular chromosome

A/Map position: circular chromosome

Query Match 70.0%; Score 35; DB 2; Length 689;

Best Local Similarity 60.0%; Pred. No. 40;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVARARVW 10  
DB 679 DVAAARVW 688

RESULT 14

Wed Sep 1 08:24:16 2004

us-09-508-095-8.aug31.rpr

Page 5

T13948  
atypical protein kinase C isotype-specific interacting protein ASIP - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C/Accession: T13948  
R/Author: Y.; Hirose, T.; Tamai, Y.; Hirai, S.; Nagashima, Y.; Fujimoto, T.; Tabuse, Y.;  
J. Cell Biol. 143, 95-106, 1998  
A/Title: An atypical PKC directly associates and colocalizes at the epithelial tight jun  
A/Reference number: Z17827; MUID:98437350; PMID:9763423  
A/Accession: T13948  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1337 <I2U>  
A/Cross-references: EMBL:AB005549; NID:93868777; PIDN:BAA34216.1; PID:93868778  
C/Genetics:  
A/Genes: aabp

Query Match 70.0%; Score 35; DB 2; Length 1337;  
Best Local Similarity 75.0%; Pred. No. 79;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 AAARVVM 10  
| | | | | : |  
Db 347 AMPARVIM 354

#### RESULT 15

JN0731  
hypothetical 14.2K protein - phage SPPI  
N/Alternate names: hypothetical protein 43  
C/Species: phage SPPI  
C/Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 11-May-2000  
C/Accession: JN0731; T42350; S21450  
R/Chai, S.; Szepan, U.; Lueder, G.; Trautner, T.A.; Alonso, J.C.  
Gene 129, 41-49, 1993  
A/Title: Sequence analysis of the left end of the Bacillus subtilis bacteriophage SPPI 9  
A/Reference number: JN0729; MUID:9338123; PMID:8335259  
A/Accession: JN0731  
A/Molecule type: DNA  
A/Residues: 1-117 <CHA>  
A/Cross-references: EMBL:X65941; NID:914843; PIDN:CAA46745.1; PID:9579085  
R/Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.  
Gene 204, 201-212, 1997  
A/Title: The complete nucleotide sequence and functional organization of Bacillus subtil  
A/Reference number: Z22137; MUID:98094274; PMID:9434185  
A/Accession: T42350  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-117 <ALO>  
A/Cross-references: EMBL:X97918; PIDN:CAA6504.1

Query Match 68.0%; Score 34; DB 2; Length 117;  
Best Local Similarity 60.0%; Pred. No. 11;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVAAARVVM 10  
| | | | | : |  
Db 101 EVAAARVVM 110

Search completed: September 1, 2004, 00:16:57  
Job time : 10.8 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2004, 23:54:44 ; Search time 3.4 Seconds

(without alignments)  
153.148 Million cell updates/sec

Title: US-09-508-095-8  
Perfect score: 50  
Sequence: 1 EVARARVW 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39.5	79.0	711	TRFL_HUMAN	P02788 homo sapien
2	36	72.0	708	TRFL_BUBBU	O77698 bubalus bub
3	36	72.0	1926	LPH_FABIT	P09649 oryctolagus
4	35	72.0	1927	LPH_HUMAN	P09648 homo sapien
5	35	70.0	1392	YB62_MYCLE	O49682 mycobacteri
6	35	70.0	1333	PAD3_MOUSE	O92440 mus musculu
7	35	70.0	1337	PAD3_RAT	O92440 rattus norv
8	34.5	69.0	695	TRFL_HORSE	O77811 equus caball
9	34	68.0	707	TRFL_MOUSE	P08071 mus musculu
10	34	68.0	855	ENV_FIVT2	O02282 feline immu
11	33	68.0	940	WAT4_SCHCO	P37938 schizophr11
12	33	66.0	285	REP2_ECOLI	P03065 escherichia
13	33	66.0	331	SYM_LISIN	O92985 listeria in
14	33	66.0	1033	YDKG_SCHPO	P87115 schizosacch
15	33	66.0	1434	VG65_HSV11	O00106 ictaluriad h
16	33	64.0	124	R35A_CABEU	P49180 caenorhadi
17	32	64.0	220	YGAE_ECOLI	P37338 escherichia
18	32	64.0	382	GAU7_TRIE	O96911 trichoderma
19	32	64.0	804	RSQ4_HUMAN	O95294 homo sapien
20	31.5	63.0	704	TRFL_PIG	P14632 sus scrofa
21	31	62.0	108	PA24_TRIST	P82895 trimeresuru
22	31	62.0	122	RT19_SCHPO	O98830 schizosacch
23	31	62.0	150	PA25_TRIST	P82896 trimeresuru
24	31	62.0	150	MOAE_HAEIN	P43308 haemophilus
25	31	62.0	185	VG66_HAEIN	P71387 haemophilus
26	31	62.0	242	YX81_STRCO	P45497 streptomyce
27	31	62.0	258	PPVK_XANCP	O89499 xanthomonas
28	31	62.0	275	TAUC_ECOLI	O47539 escherichia
29	31	62.0	313	TRBP_ECOLI	P17224 escherichia
30	31	62.0	358	YAROB_CORPS	P66749 corynebacte
31	31	62.0	406	YBDN_ECOLI	P77216 escherichia
32	31	62.0	481	MTN3_MOUSE	O35701 mus musculu

## ALIGNMENTS

RESULT 1	ID	TRFL_HUMAN	STANDARD	PRT	711 AA
AC	P02788	O00756	Q16780	Q16785	Q16789
AC	Q9H123				Q96K24
DT	21-JUL-1986	(Rel. 01, Created)			Q96K25
DT	15-JUL-1999	(Rel. 38, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Lactoferrin precursor (Lactoferrin) [Contains: Lactoferron A, Lactoferron B, Lactoferron C].				
DE	Lactoferron B, Lactoferron C].				
GN	LTF OR LTF.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid=9606;				
PN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Mammary gland;				
RX	MEDLINE=90384839; PubMed=2402455;				
RA	Rey M.W., Moloshuk S.L., de Boer H.A., Pieper F.R.,				
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.",				
RT	Nucleic Acids Res. 18:5286-5288(1990).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Mammary gland;				
RA	Cho Y.Y.;				
RN	Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.				
RP	SEQUENCE FROM N.A.				
RC	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Mammary gland;				
RA	Liang Q., Jimenez-Flores R., Richardson T.,				
RT	"Molecular cloning and sequence analysis of human lactoferrin.",				
RT	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Bone marrow;				
RA	Wei X., Han J., Rado T.A.;				
RT	"Human neutrophil lactoferrin coding and 5' flanking region DNA				
RT	sequences.",				
RT	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Mammary gland;				
RA	Cheng H., Chen X., Huan L.;				
RT	"cDNA cloning and sequence analysis of human lactoferrin.",				
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.				
RN	[7]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Prostate;				
RX	MEDLINE=22388257; PubMed=12477932;				
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.W., Schuler G.D.,				

34	31	62.0	521	1	ICBA_HUMAN	Q92851 homo sapien
35	31	62.0	521	1	TSAS_RICIS	P37917 rickettsia
36	31	62.0	560	1	J160_HORVU	O00531 hordeum vul
37	31	62.0	729	1	GLB2_XANCP	O896d1 xanthomonas
38	31	62.0	1033	1	TR1L_ECOLI	P10486 escherichia
39	31	62.0	1335	1	DE3A_DEIRA	Q91X08 deinococcus
40	31	62.0	1411	1	Y297_HUMAN	O15040 homo sapien
41	30.5	61.0	708	1	TRFL_BOVIN	P24627 bos taurus
42	30	60.0	141	1	HBA_MICGA	P41331 microcephal
43	30	60.0	159	1	RECY_RALSO	O87195 ralsstonia s
44	30	60.0	175	1	L1ST_ECOLI	P37025 escherichia
45	30	60.0	195	1	YORC_ECOLI	P33365 escherichia

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Boraldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Cammici P., Prange C.,  
 RA Rabe S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollighy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Whiting J., Helton E., Keltman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,  
 RA Butlerfield A.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [18]  
 RP SEQUENCE OF 3-711 FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=90326549; PubMed=2374734;  
 RA Powell M.J., Ogden J.E.;  
 RT "Nucleotide sequence of human lactoferrin cDNA.",  
 RL Nucleic Acids Res. 18:4013-4013(1990).  
 RN [9]  
 RP SEQUENCE OF 20-711.  
 RX MEDLINE=85076667; PubMed=6510420;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
 RA Legend D., Spik G., Montreuil J., Jolles P.;  
 RT "Human lactoferrin: amino acid sequence and structural  
 RT comparisons with other transferrins.",  
 RL Eur. J. Biochem. 145:659-666(1985).  
 RN [10]  
 RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
 RX MEDLINE=82046817; PubMed=6794640;  
 RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "The present state of the human lactoferrin sequence. Study and  
 RT alignment of the cyanogen bromide fragments and characterization of  
 RT N- and C-terminal domains.",  
 RL Biochim. Biophys. Acta 670:243-254(1981).  
 RN [11]  
 RP SEQUENCE OF 609-711.  
 RX MEDLINE=82262043; PubMed=7049727;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "An 88 amino acid long C-terminal sequence of human  
 RT lactoferrin.",  
 RL FEBS Lett. 142:107-110(1982).  
 RN [12]  
 RP SEQUENCE OF 436-711 FROM N.A.  
 RX MEDLINE=88001031; PubMed=3477300;  
 RA Rado T.A., Wei X., Benz E.J. Jr.;  
 RT "Isolation of lactoferrin cDNA from a human myeloid library and  
 RT expression of mRNA during normal and leukemic myelopoiesis.",  
 RL Blood 70:989-993(1987).  
 RN [13]  
 RP SEQUENCE OF 237-711 FROM N.A.  
 RA McComble W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
 RA Nhan M., Parrell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
 RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
 RA Chan C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
 RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
 RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
 RA Sagipanti J.L.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
 RX MEDLINE=90064528; PubMed=2585506;  
 RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
 RT "Structure of human lactoferrin: crystallographic structure analysis  
 RT and refinement at 2.8-A resolution.",

RL J. Mol. Biol. 209:711-734(1989).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RA Haridas M., Anderson B.F., Baker E.N.;  
 RT "Structure of human dimeric lactoferrin refined at 2.2-A  
 RT resolution.",  
 RL Acta Crystallogr. D 51:629-646(1995).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
 RX MEDLINE=9716796; PubMed=9003186;  
 RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
 RA Baker E.N.;  
 RT "Maturation of the histidine ligand in human lactoferrin: iron  
 RT binding properties and crystal structure of the histidine-  
 RT 253--methionine mutant.",  
 RL Biochemistry 36:341-346(1997).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=99190892; PubMed=10089347;  
 RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
 RT "Structure of recombinant human lactoferrin expressed in *Aspergillus*  
 RT *awamori*.",  
 RL Acta Crystallogr. D 55:403-407(1999).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=99192677; PubMed=10089508;  
 RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.B., Baker E.N.;  
 RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
 RT and analysis of ligand-induced conformational change.",  
 RL Acta Crystallogr. D 54:1119-1135(1998).  
 RN [19]  
 RP CHARACTERIZATION OF LACTOFERROXINS.  
 RX MEDLINE=91166929; PubMed=1359293;  
 RA Tani F., Ito K., Chiba H., Yoshikawa M.;  
 RT "Isolation and characterization of opiod antagonist peptides derived  
 RT from human lactoferrin.",  
 RL Agric. Biol. Chem. 54:1803-1810(1990).  
 RN [20]  
 RP VARIANTS THR-30 AND ARG-48.  
 RX MEDLINE=99091914; PubMed=9873069;  
 RA Kilmworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
 RA Qumsteyh W.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
 RA Sugars J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
 RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
 RA Hejmacnik J.F., Teng C.T.;  
 RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
 RT corneal dystrophy): exclusion of linkage to lactoferrin gene.",  
 RL Mol. Vision 4:31-32(1998).  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
 CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
 CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
 CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: Composed of two homologous domains.  
 CC -1- SIMILARITY: Belongs to the transferrin family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL: X53961; CAA57914.1; -  
 DR EMBL: U07643; AAB60324.1; -  
 DR EMBL: M93150; AAA6159.1; -  
 DR EMBL: M83202; AAA9511.1; -  
 DR EMBL: M83205; AAA8656.1; -





FT TURN 251 252  
 FT STRAND 255 257  
 FT HELIX 258 260  
 FT TURN 261 263  
 FT STRAND 267 270  
 FT STRAND 273 277  
 FT HELIX 283 297  
 FT TURN 299 300  
 FT TURN 307 308  
 FT TURN 312 313  
 FT TURN 321 322  
 FT STRAND 325 328  
 FT TURN 331 332  
 FT HELIX 335 351  
 FT HELIX 354 362  
 FT STRAND 364 369  
 FT HELIX 371 384  
 FT TURN 385 385  
 FT STRAND 388 393  
 FT HELIX 396 404  
 FT TURN 405 406  
 FT STRAND 410 413  
 FT HELIX 415 422  
 FT TURN 423 425  
 FT STRAND 426 434  
 FT TURN 438 441  
 FT TURN 444 446  
 FT STRAND 452 459  
 FT TURN 460 461  
 FT TURN 463 464  
 FT HELIX 467 469  
 FT TURN 471 472  
 FT STRAND 474 477  
 FT TURN 480 481  
 FT TURN 483 486  
 FT HELIX 487 497  
 FT TURN 503 505  
 FT STRAND 508 510  
 FT TURN 512 513  
 FT TURN 516 517  
 FT TURN 519 520  
 FT TURN 522 523  
 FT TURN 528 529  
 FT TURN 533 534  
 FT TURN 538 539  
 FT TURN 541 542  
 FT HELIX 544 553  
 FT STRAND 554 555  
 FT STRAND 559 563  
 FT HELIX 564 569  
 FT TURN 570 570  
 FT HELIX 578 581  
 FT TURN 582 582  
 FT TURN 585 586  
 FT STRAND 588 591  
 FT TURN 593 594  
 FT STRAND 597 599  
 FT TURN 600 601  
 FT TURN 603 605  
 FT STRAND 609 611  
 FT STRAND 615 619  
 FT HELIX 620 637  
 FT TURN 639 640  
 FT TURN 642 647  
 FT TURN 650 651  
 FT TURN 654 655

Query Match 72.0%; Score 36; DB 1; Length 708;  
 Best Local Similarity 87.5%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

3 AARAYVW 10  
 |||||

DE 359 AARAYVW 366  
 RESULT 3  
 LPH\_RABIT STANDARD; PRT, 1926 AA.  
 ID\_LPH\_RABIT  
 AC P09849;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, last sequence update)  
 DT 30-MAY-2000 (Rel. 39, last annotation update)  
 DE Lactase-phlorizin hydrolase precursor (lactase-glycosylceramidase)  
 DE [Includes: lactase (EC 3.2.1.108); Phlorizin hydrolase (EC 3.2.1.62)].  
 GN LCT OR LPH.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_Taxid=9986;  
 [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=New Zealand white;  
 RX MEDLINE=89030634; PubMed=2460343;  
 RA Mantel N., Villa M., Ezler T., Wacker H., Boll W., James P.,  
 RA Hunziker W., Semenza G.;  
 RT "Complete primary structure of human and rabbit lactase-phlorizin  
 RT hydrolase: implications for biosynthesis, membrane anchoring and  
 RT evolution of the enzyme.";  
 RL EMBO J. 7:2705-2713 (1988).  
 [2]  
 RP ACTIVE SITES.  
 RX MEDLINE=92406791; PubMed=1388157;  
 RA Wacker H., Keller P., Falchetto R., Legler G., Semenza G.;  
 RT "Location of the two catalytic sites in intestinal lactase-phlorizin  
 RT hydrolase. Comparison with sucrose-isomaltase and with other  
 RT glycosidases, the membrane anchor of lactase-phlorizin hydrolase.";  
 RL J Biol. Chem. 267:18744-18752 (1992).  
 CC -1- FUNCTION: LPH splits lactose in the small intestine.  
 CC -1- CATALYTIC ACTIVITY: Lactose + H(2)O = D-glucose + D-galactose.  
 CC -1- CATALYTIC ACTIVITY: Glycosyl-N-acetylphingosine + H(2)O = a sugar +  
 CC -1- N-acetylphingosine.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Brush border.  
 CC -1- DOMAIN: THE SEQUENCE EXHIBITS 4 REGIONS (I-IV) OF INTERNAL  
 CC -1- HOMOLOG: THEREFORE LPH MIGHT HAVE EVOLVED BY TWO CYCLES OF  
 CC PARTIAL GENE DUPLICATION.  
 CC -1- SIMILARITY: Belongs to family 1 of glycosyl hydrolases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X07995; CAA30802.1; -  
 DR PIR; S01169; S01169.  
 DR HSSP; P26205; ICBG.  
 DR InterPro; IPR001360; Glyco\_hydro\_1.  
 DR Pfam; PF00232; Glyco\_hydro\_1; 5.  
 DR PRINTS; PR00131; GLHYDRLASE1.  
 DR ProDom; PD000650; Glyco\_hydro\_1; 4.  
 DR PROSITE; PS00572; GLYCOSYL\_HYDROL\_F1\_1; 2.  
 DR PROSITE; PS00653; GLYCOSYL\_HYDROL\_F1\_2; 3.  
 DR HydroLase; Glycosidase; Zymogen; Signal; Transmembrane; Repeat.  
 FT SIGNAL 1 19  
 FT PROPEP 20 866  
 FT CHAIN 867 1926  
 FT DOMAIN 20 1882  
 FT TRANSMEM 1883 1901  
 FT DOMAIN 1902 1926  
 FT DOMAIN 79 1800  
 FT REPEAT 79 172  
 FT REPEAT 360 845  
 FT REPEAT 2.

```

FT REPEAT 981 1367 3.
FT REPEAT 1375 1800 4.
FT ACT_SITE 1063 1063 PROTON DONOR (POTENTIAL).
FT ACT_SITE 1271 1271 NUCLEOPHILE.
FT ACT_SITE 1536 1536 PROTON DONOR (POTENTIAL).
FT ACT_SITE 1747 1747 NUCLEOPHILE.
SQ SEQUENCE 1926 AA; 217847 MW; 2A21A737D0CC7A CRC64;

Query Match
Best Local Similarity 72.0%; Score 36; DB 1; Length 1927;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Cy 1 EVARARVVM 10
Db 673 EVPSKATVM 882

RESULT 4
LPH_HUMAN STANDARD; PRT; 1927 AA.
AC P09848;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lactase-phlorizin hydrolase precursor (lactase-glycosylceramidase)
DE [includes: Lactase (EC 3.2.1.108), Phlorizin hydrolase (EC 3.2.1.62)].
GN LCT OR LPH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9030634; PubMed=2460343;
RA Mantel N., Villa M., Enzler T., Wacker H., Boll W., James P.,
RA Hunziker W., Semenza G.;
RT "Complete primary structure of human and rabbit lactase-phlorizin
RT hydrolase: implications for biosynthesis, membrane anchoring and
RT evolution of the enzyme."
RL EMBO J. 7:2705-2713(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91206402; PubMed=1902057;
RA Boll W., Wagner P., Mantel N.;
RT "Structure of the chromosomal gene and cDNAs coding for lactase-
RT phlorizin hydrolase in humans with adult-type hypolactasia or
RT persistence of lactase."
RL Am. J. Hum. Genet. 48:889-902(1991).
CC -1- FUNCTION: LPH splits lactose in the small intestine.
CC -1- CATALYTIC ACTIVITY: Lactose + H(2)O = D-glucose + D-galactose.
CC -1- CATALYTIC ACTIVITY: Glycosyl-N-acetylphosphoglycine + H(2)O = a sugar +
CC N-acetylphosphoglycine.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Brush border.
CC -1- TISSUE SPECIFICITY: Intestine.
CC -1- DOMAIN: THE SEQUENCE EXHIBITS 4 REGIONS (I-IV) OF INTERNAL
CC HOMOLOG; THEREFORE LPH MIGHT HAVE EVOLVED BY TWO CYCLES OF
CC PARTIAL GENE DUPLICATION.
CC -1- DISEASE: Defects in the upstream vicinity of LCT are the cause of
CC disaccharide intolerance III [MIM:223100]; also known as adult
CC lactase deficiency or adult-type hypolactasia. In many human
CC populations the activity of LCT declines in adults, leading to
CC adult-type hypolactasia, whereas in other populations the high
CC activity persists.
CC -1- SIMILARITY: Belongs to family 1 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; X07994; CAA30801.1; -.
DR EMBL; M61850; AAA59504.1; -.
DR EMBL; M61834; AAA59504.1; JOINED.
DR EMBL; M61835; AAA59504.1; JOINED.
DR EMBL; M61836; AAA59504.1; JOINED.
DR EMBL; M61837; AAA59504.1; JOINED.
DR EMBL; M61838; AAA59504.1; JOINED.
DR EMBL; M61839; AAA59504.1; JOINED.
DR EMBL; M61840; AAA59504.1; JOINED.
DR EMBL; M61841; AAA59504.1; JOINED.
DR EMBL; M61842; AAA59504.1; JOINED.
DR EMBL; M61843; AAA59504.1; JOINED.
DR EMBL; M61844; AAA59504.1; JOINED.
DR EMBL; M61845; AAA59504.1; JOINED.
DR EMBL; M61846; AAA59504.1; JOINED.
DR EMBL; M61847; AAA59504.1; JOINED.
DR EMBL; M61848; AAA59504.1; JOINED.
DR EMBL; M61849; AAA59504.1; JOINED.
DR F1R; S01168; S01168.
DR HSP; P26205; ICBG.
DR Genew; HGNC:6530; LCT.
DR MIM; 603202; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0000016; F:lactase activity; TAS.
DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00232; Glyco_hydro_1; 4.
DR ProDom; PD00650; Glyco_hydro_1; 4.
DR PRINTS; PR00131; GLYHDRLASE1.
DR ProDom; PD00650; Glyco_hydro_1; 4.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_2; 3.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_3; 3.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_4; 4.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_5; 5.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_6; 6.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_7; 7.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_8; 8.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_9; 9.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_10; 10.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_11; 11.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_12; 12.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_13; 13.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_14; 14.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_15; 15.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_16; 16.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_17; 17.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_18; 18.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_19; 19.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_20; 20.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_21; 21.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_22; 22.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_23; 23.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_24; 24.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_25; 25.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_26; 26.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_27; 27.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_28; 28.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_29; 29.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_30; 30.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_31; 31.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_32; 32.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_33; 33.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_34; 34.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_35; 35.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_36; 36.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_37; 37.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_38; 38.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_39; 39.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_40; 40.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_41; 41.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_42; 42.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_43; 43.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_44; 44.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_45; 45.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_46; 46.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_47; 47.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_48; 48.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_49; 49.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_50; 50.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_51; 51.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_52; 52.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_53; 53.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_54; 54.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_55; 55.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_56; 56.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_57; 57.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_58; 58.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_59; 59.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_60; 60.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_61; 61.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_62; 62.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_63; 63.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_64; 64.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_65; 65.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_66; 66.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_67; 67.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_68; 68.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_69; 69.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_70; 70.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_71; 71.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_72; 72.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_73; 73.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_74; 74.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_75; 75.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_76; 76.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_77; 77.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_78; 78.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_79; 79.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_80; 80.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_81; 81.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_82; 82.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_83; 83.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_84; 84.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_85; 85.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_86; 86.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_87; 87.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_88; 88.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_89; 89.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_90; 90.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_91; 91.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_92; 92.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_93; 93.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_94; 94.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_95; 95.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_96; 96.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_97; 97.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_98; 98.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_99; 99.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_100; 100.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_101; 101.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_102; 102.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_103; 103.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_104; 104.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_105; 105.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_106; 106.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_107; 107.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_108; 108.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_109; 109.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_110; 110.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_111; 111.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_112; 112.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_113; 113.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_114; 114.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_115; 115.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_116; 116.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_117; 117.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_118; 118.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_119; 119.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_120; 120.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_121; 121.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_122; 122.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_123; 123.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_124; 124.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_125; 125.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_126; 126.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_127; 127.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_128; 128.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_129; 129.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_130; 130.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_131; 131.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_132; 132.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_133; 133.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_134; 134.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_135; 135.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_136; 136.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_137; 137.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_138; 138.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_139; 139.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_140; 140.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_141; 141.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_142; 142.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_143; 143.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_144; 144.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_145; 145.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_146; 146.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_147; 147.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_148; 148.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_149; 149.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_150; 150.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_151; 151.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_152; 152.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_153; 153.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_154; 154.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_155; 155.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_156; 156.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_157; 157.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_158; 158.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_159; 159.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_160; 160.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_161; 161.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_162; 162.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_163; 163.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_164; 164.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_165; 165.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_166; 166.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_167; 167.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_168; 168.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_169; 169.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_170; 170.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_171; 171.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_172; 172.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_173; 173.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_174; 174.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_175; 175.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_176; 176.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_177; 177.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_178; 178.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_179; 179.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_180; 180.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_181; 181.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_182; 182.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_183; 183.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_184; 184.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_185; 185.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_186; 186.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_187; 187.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_188; 188.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_189; 189.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_190; 190.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_191; 191.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_192; 192.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_193; 193.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_194; 194.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_195; 195.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_196; 196.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_197; 197.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_198; 198.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_199; 199.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_200; 200.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_201; 201.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_202; 202.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_203; 203.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_204; 204.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_205; 205.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_206; 206.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_207; 207.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_208; 208.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_209; 209.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_210; 210.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_211; 211.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_212; 212.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_213; 213.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_214; 214.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_215; 215.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_216; 216.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_217; 217.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_218; 218.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_219; 219.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_220; 220.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_221; 221.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_222; 222.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_223; 223.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_224; 224.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_225; 225.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_226; 226.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_227; 227.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_228; 228.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_229; 229.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_230; 230.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_231; 231.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_232; 232.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_233; 233.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_234; 234.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_235; 235.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_236; 236.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_237; 237.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_238; 238.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_239; 239.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_240; 240.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_241; 241.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_242; 242.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_243; 243.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_244; 244.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_245; 245.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_246; 246.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_247; 247.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_248; 248.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_249; 249.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_250; 250.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_251; 251.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_252; 252.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_253; 253.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_254; 254.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_255; 255.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_256; 256.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_257; 257.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_258; 258.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_259; 259.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_260; 260.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_261; 261.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_262; 262.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_263; 263.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_264; 264.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_265; 265.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_266; 266.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_267; 267.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_268; 268.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_269; 269.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_270; 270.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_271; 271.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_272; 272.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_273; 273.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_274; 274.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_275; 275.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_276; 276.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_277; 277.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_278; 278.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_279; 279.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_280; 280.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_281; 281.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_282; 282.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_283; 283.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_284; 284.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_285; 285.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_286; 286.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_287; 287.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_288; 288.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_289; 289.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_290; 290.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_291; 291.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_292; 292.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_293; 293.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_294; 294.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_295; 295.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_296; 296.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_297; 297.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_298; 298.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_299; 299.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_300; 300.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_301; 301.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_302; 302.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_303; 303.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_304; 304.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_305; 305.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_306; 306.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_307; 307.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_308; 308.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_309; 309.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_310; 310.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_311; 311.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_312; 312.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_313; 313.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_314; 314.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_315; 315.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_316; 316.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_317; 317.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_318; 318.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_319; 319.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_320; 320.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_321; 321.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_322; 322.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_323; 323.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_324; 324.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_325; 325.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_326; 326.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_327; 327.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_328; 328.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_329; 329.
DR PROSITE; PS00573; GLYCOSYL_HYDROL_F1_330; 330.

```

OK NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Smith D.R., Robinson K.;  
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA STRAIN=TN;  
RX MEDLINE=21128732; PubMed=11234002;  
RA Cole S.T., Biglmeier K., Parkhill J., James K.D., Thomson N.R.,  
RA Wheeler P.R., Honore N., Garnier T., Churruarin C., Harris D.,  
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,  
RA Holtz S., Hornby K., Jagels K., Lacroix C., Maclean J., Mable S.,  
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
RA Rutter S., Seeger K., Simon S., Simmonds M., Skilton J., Squares R.,  
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
RA Barrett B.G.;  
RL "Massive gene decay in the leprosy bacillus."  
RT Nature 409:1007-1011 (2001).  
CC -1- SIMILARITY: Belongs to the UPF0051 (ycf24) family.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL: U00013; AAA17119.1; -  
DR EMBL: 299125; CAB16170.1; -  
DR EMBL: AL583919; CAC30102.1; -  
DR PIR: S72753; S72753.  
DR Leprosoma: ML0594; -  
DR InterPro: IPR000825; UPF0051.  
DR Pfam: PF01458; UPF0051.1.  
DR Hypothetical protein: Complete proteome.  
SQ SEQUENCE 392 AA; 42202 MW; 891162F7CA494C6A CRC64;  
Query Match 70.0%; Score 35; DB 1; Length 392;  
Best Local Similarity 70.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Cy 1 EVAPARVW 10  
Db 164 EAARLTVM 173  
RESULT 6  
RAD3\_MOUSE STANDARD; PRT; 1333 AA.  
ID RAD3\_MOUSE  
AC Q99NH2;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Partitioning defective 3 homolog (PAR-3) (PAR-3) (Atypical PKC  
DE isotype-specific interacting protein) (ASIP) (Ephrin interacting  
DE protein) (PPIP).  
GN PAR3 OR PAR3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RA STRAIN=NIH Swiss; TISSUE=Embryo;  
RX MEDLINE=9912117; PubMed=920925;  
RA Lin D., Gish G.D., Songyang Z., Pawson T.;  
RT "The carboxyl terminus of B class ephrins constitutes a PDZ domain  
RT binding motif."  
RL J. Biol. Chem. 274:3726-3733 (1999).  
RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), SUBCELLULAR LOCATION,  
RP PHOSPHORYLATION BY PRKCZ, INTERACTION WITH PRKCI AND PAR6A, SUBUNIT  
RP OF A COMPLEX CONTAINING PAR6A AND CDC42, AND KOTAEINESIS OF  
RP 824-SER-826.  
RC STRAIN=NIH Swiss;  
RX MEDLINE=20394297; PubMed=10934475;  
RA Lin D., Edwards A.S., Fawcett J.P., Wdamsu G., Scott J.D., Pawson T.;  
RT "A mammalian PAR-3-PAR-6 complex implicated in Cdc42/Rac1 and APC  
RT signalling and cell polarity."  
RL Nat. Cell Biol. 2:540-547 (2000).  
RN [3]  
RP INTERACTION WITH PAR6B, AND SUBUNIT OF A COMPLEX CONTAINING PAR6B;  
RP PRKCI AND CDC42.  
RX MEDLINE=20394296; PubMed=10934474;  
RA Jobery G., Petersen C., Gao L., Macara I.G.;  
RT "The cell-polarity protein Par6 links Par3 and atypical protein kinase  
RT C to Cdc42."  
RL Nat. Cell Biol. 2:531-539 (2000).  
RN [4]  
RP INTERACTION WITH F1LR.  
RX MEDLINE=21340266; PubMed=11447115;  
RA Ebner K., Suzuki A., Horikoshi Y., Hirose T.,  
RA Meyer zu Brucke M.-K., Ono S., Vestweber D.;  
RT "The cell polarity protein ASIP/Par-3 directly associates with  
RT junctional adhesion molecule (JAM)."  
RL EMBO J. 20:3738-3748 (2001).  
RN [5]  
RP INTERACTION WITH F1LR AND PAR6B.  
RX MEDLINE=2128709; PubMed=11839275;  
RA Gao L., Jobery G., Macara I.G.;  
RT "Assembly of epithelial tight junctions is negatively regulated by  
RT Par6."  
RL Curr. Biol. 12:221-225 (2002).  
CC -1- FUNCTION: Adapter protein involved in asymmetrical cell division  
CC and cell polarization processes. Plays a role in the formation of  
CC epithelial tight junctions. Association with PAR6B may prevent  
CC the interaction of PAR3 with F1LR/JAM1, thereby preventing tight  
CC junction assembly. The PAR6-PAR3 complex links GTP-bound Rho  
CC small GTPases to atypical protein kinase C proteins.  
CC -1- SUBUNIT: Interacts with PAR6A, PAR6B, PRKCI and CDC42. Part of a  
CC complex with PAR6A or PAR6B, PRKCI or PRKCZ and CDC42 or RAC1.  
CC -1- INTERACTS WITH F1LR/JAM1.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; membrane associated. Localized  
CC along the cell-cell contact region. Colocalizes with PAR6A and  
CC PRKCI at epithelial tight junctions.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event: alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=180 kDa;  
CC IsoId=Q99NH2-1; Sequence=Displayed;  
CC Name=2; Synonyms=150 kDa;  
CC IsoId=Q99NH2-2; Sequence=VSP\_007474;  
CC Name=3; Synonyms=100 kDa;  
CC IsoId=Q99NH2-3; Sequence=VSP\_007473;  
CC -1- TISSUE SPECIFICITY: All isoforms are expressed in heart, while  
CC expression in brain is mainly limited to isoform 1, and to isoform  
CC 3 to a weaker level.  
CC -1- DEVELOPMENTAL STAGE: Isoforms 1 and isoform 3 are expressed from E  
CC 9.5 to E14.5, while isoform 2 is not expressed.  
CC -1- PTM: Phosphorylated by PRKCZ.  
CC -1- SIMILARITY: Belongs to the PAR3 family.  
CC -1- SIMILARITY: Contains 3 PDZ/DHR domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AY026057; AAK07669.1; -  
DR HSP: Q12923; 3PDZ.  
DR MGI:2135608; Pard3.

DR GO: GO:0005913; C:cell-cell adherens junction; IDA.  
 DR GO: GO:0005515; P:protein binding; IPI.  
 DR GO: GO:0016337; P:cell-cell adhesion; IC.  
 DR InterPro: IPR001478; PDZ.  
 DR Pfam: PF00595; PDZ: 3.  
 DR SMART: SMO0228; PDZ: 3.  
 DR PROSITE: PS5106; PDZ: 3.  
 KW Cell cycle; Cell division; Tight junction; Membrane; Repeat;  
 KW Coiled coil; Phosphorylation; Alternative splicing.  
 FT DOMAIN 271 359  
 FT DOMAIN 461 546  
 FT DOMAIN 590 677  
 FT DOMAIN 712 936  
 FT DOMAIN 980 1038  
 FT DOMAIN 1046 1078  
 FT DOMAIN 1145 1168  
 FT DOMAIN 1195 1218  
 FT DOMAIN 1274 1295  
 FT MOD RES 824 824  
 FT VARSPPLIC 741 744  
 FT VARSPPLIC 745 1333  
 FT VARSPPLIC 1030 1333  
 FT MOTAGEN 824 826  
 FT  
 FT  
 SQ SEQUENCE 1333 AA; 149060 MW; A67825C6DCE86 CRC64;  
 Query Match 70.0%; Score 35; DB 1; Length 1333;  
 Best Local Similarity 75.0%; Pred. No. 43;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AARAVVM 10  
 DB 347 AMRAVIM 354

RESULT 7  
 PAD\_RAT  
 ID PAD3\_RAT STANDARD; PRT; 1337 AA.  
 AC Q92340;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Partitioning-defective 3 homolog (PAR-3) (PAR-3) (Atypical PKC  
 DE Isozyme-specific interacting protein) (ASIP) (Atypical PKC specific  
 DE Binding protein) (ASBP).  
 GN PAR3 OR PAR3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH PRKCZ AND PRKCZ.  
 RP TISSUE=Fibroblast;  
 RX MEDLINE=98437350; PubMed=9763423;  
 RX Izumi Y., Hirose T., Tamai Y., Hirai S.-I., Nagashima Y., Fujimoto T.,  
 RA Tabuse Y., Kempner K.J., Ohno S.,  
 RT "An atypical PKC directly associates and colocalizes at the epithelial  
 RT tight junction with ASIP, a mammalian homologue of caenorhabditis  
 RT elegans polarity protein PAR-3.";  
 RT J. Cell Biol. 143:95-106(1998).  
 RL [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 2), SUBCELLULAR LOCATION, AND  
 RP PHOSPHORYLATION OF SER-827.  
 RX MEDLINE=22040480; PubMed=12045219;  
 RX Hirose T., Izumi Y., Nagashima Y., Tamai-Nagai Y., Kurihara H.,  
 RA Sakai T., Suzuki Y., Yamataka T., Suzuki A., Mizuno K., Ohno S.,  
 RT "Involvement of ASIP/PAR-3 in the promotion of epithelial tight  
 RT junction formation.";  
 RT J. Cell Sci. 115:2485-2495(2002).

CC -1- FUNCTION: Adapter protein involved in asymmetrical cell division  
 CC and cell polarization processes. Seems to play a central role in  
 CC the formation of epithelial tight junctions. Association with  
 CC PAR6B may prevent the interaction of PAR3 with FilR/JM1,  
 CC thereby preventing tight junction assembly. The PAR3-PAR6B  
 CC complex links GTP-bound Rho small GTPases to atypical protein  
 CC kinase C proteins.  
 CC -1- SUBUNIT: Interacts with PAR6A, PAR6B and FilR/JM1 via its PDZ 1  
 CC domain (By similarity). Interacts with PRKC1. Interacts with PRKCZ  
 CC (Probable). Part of a complex with PAR6A or PAR6B, PRKC1 or  
 CC PRKCZ and Cdc42 or Rac1 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; membrane associated. Localized  
 CC along the cell-cell contact region. Colocalizes with PRKCZ at the  
 CC apical edge of tight junctions.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=180 kDa;  
 CC Name=2; Synonyms=150 kDa;  
 CC IsoId=Q92340-2; Sequence=VSP\_007475;  
 CC -1- TISSUE SPECIFICITY: Isoform 1 is predominantly expressed in lung,  
 CC glandular stomach, prostate, ovary and uterus. Isoform 1 is also  
 CC expressed in brain, with a high expression in the cortex,  
 CC hippocampus and in the striatum. Isoform 2 is predominantly  
 CC expressed in intestinal epithelial cells, kidney and prostate.  
 CC -1- PTM: Phosphorylated by PRKCZ (By similarity). The phosphorylated  
 CC form is concentrated at the most apical tip of cell-cell contacts  
 CC during the initial phase of tight junction formation.  
 CC -1- SIMILARITY: Belongs to the PAR3 family.  
 CC -1- SIMILARITY: Contains 3 PDZ/DHR domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL; AB005549; BAA34216.1; -  
 CC DR PIR; T13948; T13948.  
 CC DR HSSP; Q12923; 3PDZ.  
 CC DR InterPro; IPR001478; PDZ.  
 CC DR Pfam; PF00595; PDZ: 3.  
 CC DR SMART; SMO0228; PDZ: 3.  
 CC DR PROSITE; PS5106; PDZ: 3.  
 KW Cell cycle; Cell division; Tight junction; Membrane; Repeat;  
 KW Coiled coil; Phosphorylation; Alternative splicing.  
 FT DOMAIN 271 359  
 FT DOMAIN 461 546  
 FT DOMAIN 590 677  
 FT DOMAIN 712 936  
 FT DOMAIN 984 1042  
 FT DOMAIN 1050 1082  
 FT DOMAIN 1149 1172  
 FT DOMAIN 1199 1222  
 FT DOMAIN 1278 1299  
 FT MOD RES 827 827  
 FT VARSPPLIC 1034 1337  
 FT  
 SQ SEQUENCE 1337 AA; 149448 MW; EC980C510E85299C CRC64;  
 Query Match 70.0%; Score 35; DB 1; Length 1337;  
 Best Local Similarity 75.0%; Pred. No. 43;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AARAVVM 10  
 DB 347 AMRAVIM 354

RESULT 8  
 TRFL\_HORSE

ID TRFL\_HORSE STANDARD; PRT; 695 AA.  
AC 077811;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Lactotransferrin precursor (Lactoferrin) (Fragment).  
GN LTF.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_Taxid=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Paramesivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
RT "cDNA sequence of mare lactoferrin."  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
RC TISSUE=Milk;  
RX MEDLINE=99296631; PubMed=10366507;  
RA Sharma A.K., Paramesivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
RT "Three-dimensional structure of mare dimeric lactoferrin at 2.6-A resolution."  
RL J. Mol. Biol. 289:303-317(1999).  
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING OF AN ANION, USUALLY BICARBONATE.  
CC -1- SUBUNIT: Monomer.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- DOMAIN: Composed of two homologous domains.  
CC -1- SIMILARITY: Belongs to the transferrin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: A1010930; CA09407.1; -.  
DR PDB; 1BLX; 02-DEC-98.  
DR PDB; 1B70; 02-FEB-99.  
DR PDB; 1B7Z; 02-FEB-99.  
DR PDB; 1F9B; 10-FEB-01.  
DR PDB; 1I6B; 13-FEB-02.  
DR PDB; 1QW; 14-JAN-00.  
DR InterPro: IPR001156; Transferrin.  
DR Pfam: PF00405; Transferrin; 2.  
DR PRINTS; PRO0422; TRANSFERRIN.  
DR SMART; SM00094; TR\_FER; 2.  
DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE; PS00207; TRANSFERRIN\_3; 1.  
KM Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
FT Signal; 3d-structure.  
FT NON\_TER 1  
FT SIGNAL 1  
FT CHAIN 7 695  
FT REPEAT 7 350 LACTOTRANSFERRIN.  
FT REPEAT 351 695 1.  
FT DISULFID 15 51 2.  
FT DISULFID 25 42  
FT DISULFID 121 204  
FT DISULFID 163 179  
FT DISULFID 166 189  
FT DISULFID 176 187  
FT DISULFID 237 251  
FT DISULFID 354 386  
FT DISULFID 364 377  
FT DISULFID 411 690  
FT DISULFID 431 653  
FT DISULFID 463 538

FT DISULFID 487 681  
FT DISULFID 497 511  
FT DISULFID 508 521  
FT DISULFID 579 593  
FT DISULFID 631 636  
FT METAL 66 66  
FT METAL 98 98  
FT METAL 198 198  
FT METAL 259 259  
FT METAL 401 401  
FT METAL 439 439  
FT METAL 532 532  
FT METAL 601 601  
FT BINDING 123 123  
FT BINDING 127 127  
FT BINDING 129 129  
FT BINDING 130 130  
FT BINDING 130 130  
FT BINDING 465 465  
FT BINDING 469 469  
FT BINDING 471 471  
FT BINDING 472 472  
FT CARBOHYD 143 143  
FT CARBOHYD 287 287  
FT CARBOHYD 482 482  
FT STRAND 12 16  
FT HELIX 19 34  
FT TURN 35 36  
FT STRAND 40 44  
FT HELIX 48 56  
FT TURN 57 58  
FT STRAND 62 65  
FT HELIX 67 74  
FT TURN 76 78  
FT STRAND 80 88  
FT STRAND 95 95  
FT STRAND 97 105  
FT TURN 106 107  
FT HELIX 112 114  
FT TURN 116 117  
FT STRAND 119 122  
FT TURN 125 126  
FT TURN 128 131  
FT HELIX 132 137  
FT HELIX 139 142  
FT TURN 147 148  
FT HELIX 151 158  
FT STRAND 161 163  
FT TURN 165 166  
FT TURN 169 171  
FT HELIX 173 175  
FT TURN 176 178  
FT TURN 183 187  
FT TURN 191 192  
FT TURN 194 195  
FT HELIX 197 206  
FT TURN 207 208  
FT STRAND 212 216  
FT HELIX 219 223  
FT HELIX 227 230  
FT TURN 231 232  
FT STRAND 233 237  
FT TURN 238 240  
FT STRAND 241 244  
FT HELIX 245 247  
FT TURN 248 250  
FT STRAND 254 257  
FT STRAND 260 264  
FT HELIX 270 284  
FT TURN 286 287  
FT TURN 294 295  
FT TURN 308 309  
FT STRAND 312 315  
FT HELIX 322 326

IRON 1.  
IRON 1.  
IRON 1.  
IRON 2.  
IRON 2.  
IRON 2.  
CARBONATE 1.  
CARBONATE 1.  
CARBONATE 1.  
CARBONATE 1 (VIA AMIDE NITROGEN).  
CARBONATE 2.  
CARBONATE 2.  
CARBONATE 2 (VIA AMIDE NITROGEN).  
CARBONATE 2 (VIA AMIDE NITROGEN).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT HELIX 328 336
FT TURN 337 338
FT HELIX 341 349
FT STRAND 351 356
FT HELIX 358 370
FT TURN 371 372
FT STRAND 375 380
FT HELIX 383 392
FT TURN 393 393
FT STRAND 397 400
FT HELIX 402 410
FT TURN 411 412
FT STRAND 414 421
FT HELIX 429 433
FT STRAND 439 446
FT HELIX 454 456
FT TURN 458 459
FT STRAND 462 464
FT TURN 467 468
FT TURN 470 473
FT HELIX 474 484
FT HELIX 489 491
FT TURN 492 492
FT STRAND 496 497
FT TURN 499 500
FT TURN 503 504
FT TURN 506 507
FT STRAND 508 508
FT TURN 515 516
FT TURN 520 521
FT TURN 525 526
FT TURN 528 529
FT HELIX 531 540
FT TURN 541 542
FT STRAND 546 550
FT HELIX 551 555
FT TURN 556 557
FT TURN 559 560
FT HELIX 565 568
FT TURN 569 569
FT TURN 572 573
FT STRAND 575 578
FT TURN 580 581
FT STRAND 584 586
FT HELIX 587 592
FT STRAND 596 598
FT STRAND 602 605
FT TURN 607 609
FT HELIX 610 624
FT TURN 626 627
FT TURN 629 630
FT HELIX 631 634
FT TURN 637 638
FT TURN 641 642

Query Match 69.0%; Score 34.5; DB 1; Length 695;
Best Local Similarity 81.8%; Pred. No. 29;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

CY 1 EVAA-RARVW 10
DB 343 EVAA-RARVW 353

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA MEDLINE=87280033; PubMed=3611056;
RT Penicost B.T., Teng C.T.;
RT "Lactoferrin is the major estrogen inducible protein of mouse
RT uterine secretions.";
RN J. Biol. Chem. 262:10134-10139(1987).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Morishita K.;
RN Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin J.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Mulhally S.J.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullen P.H.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RP SEQUENCE OF 1-14 FROM N.A.
RX MEDLINE=92042099; PubMed=1939212;
RA Liu Y., Teng C.T.;
RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";
CC J. Biol. Chem. 266:21880-21885(1991).
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: Composed of two homologous domains.
CC -1- SIMILARITY: Belongs to the transferrin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03298; AAA0525.1; -
DR EMBL: D88510; BAA1363.1; -
DR EMBL: BC006904; AAH06904.1; -
DR EMBL: W74778; AAA39427.1; -
DR HSSP: P02788; 1CB6.
DR MGD: MGI:96837; Ltf.
DR InterPro: IPR001156; Transferrin.
DR Pfam: PF00405; transferrin.2.
DR PRINTS: PR00442; TRANSFERRIN.
DR SMART: SM00094; TR_FER; 2.
DR PROSITE: PS00205; TRANSFERRIN_1; 1.

```

```

DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW Signal.
FT CHAIN 1 19 BY SIMILARITY.
FT CHAIN 20 707 LACTOTRANSFERRIN.
FT REPEAT 20 357 1.
FT REPEAT 358 707 2.
FT DISULFID 27 63 BY SIMILARITY.
FT DISULFID 37 54 BY SIMILARITY.
FT DISULFID 133 216 BY SIMILARITY.
FT DISULFID 175 191 BY SIMILARITY.
FT DISULFID 188 199 BY SIMILARITY.
FT DISULFID 249 263 BY SIMILARITY.
FT DISULFID 366 398 BY SIMILARITY.
FT DISULFID 376 389 BY SIMILARITY.
FT DISULFID 423 702 BY SIMILARITY.
FT DISULFID 443 665 BY SIMILARITY.
FT DISULFID 475 550 BY SIMILARITY.
FT DISULFID 499 693 BY SIMILARITY.
FT DISULFID 509 523 BY SIMILARITY.
FT DISULFID 520 533 BY SIMILARITY.
FT DISULFID 591 605 BY SIMILARITY.
FT DISULFID 643 648 BY SIMILARITY.
FT METAL 78 78 BY SIMILARITY.
FT METAL 110 110 IRON 1 (BY SIMILARITY).
FT METAL 210 210 IRON 1 (BY SIMILARITY).
FT METAL 271 271 IRON 1 (BY SIMILARITY).
FT METAL 413 413 IRON 2 (BY SIMILARITY).
FT METAL 451 451 IRON 2 (BY SIMILARITY).
FT METAL 544 544 IRON 2 (BY SIMILARITY).
FT METAL 613 613 IRON 2 (BY SIMILARITY).
FT BINDING 135 135 CARBONATE 1 (BY SIMILARITY).
FT BINDING 139 139 CARBONATE 1 (BY SIMILARITY).
FT BINDING 141 141 CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY).
FT BINDING 142 142 CARBONATE 1 (VIA AMIDE NITROGEN) (BY SIMILARITY).
FT BINDING 477 477 CARBONATE 2 (BY SIMILARITY).
FT BINDING 481 481 CARBONATE 2 (BY SIMILARITY).
FT BINDING 483 483 CARBONATE 2 (VIA AMIDE NITROGEN) (BY SIMILARITY).
FT BINDING 484 484 CARBONATE 2 (VIA AMIDE NITROGEN) (BY SIMILARITY).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 1 2 MR -> IQG (IN REF. 1).
FT CONFLICT 25 25 R -> Q (IN REF. 2).
FT CONFLICT 82 82 S -> T (IN REF. 1).
FT CONFLICT 359 359 S -> T (IN REF. 2).
FT CONFLICT 382 382 A -> D (IN REF. 1).
FT CONFLICT 445 445 E -> G (IN REF. 2).
FT CONFLICT 629 629 L -> V (IN REF. 1).
SQ SEQUENCE 707 AA; 77865 MW; F26AE0340A4C19A8 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 707;
Best Local Similarity 55.6%; Pred. No. 37;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

GN ENV.
OS Feline immunodeficiency virus (isolate TM2) (FIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=31676;
RP SEQUENCE FROM N.A.
RX MEDLINE=91303716; PubMed=149349;
RA Kiyomatsu T., Miyazawa T., Furuya T., Shibata R., Sakai H.,
RA Sakuragi J.I., Fukasawa M., Maki N., Hasegawa A., Mikami T.,
RA Adachi A.
RT "Identification of feline immunodeficiency virus rev gene activity."
RL J. Virol. 65:4539-4542(1991).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=92198230; PubMed=1312825;
RA Maki N., Miyazawa T., Fukasawa M., Hasegawa A., Hayami M., Mikami T.,
RA Mikami T.
RT "Molecular characterization and heterogeneity of feline
RT immunodeficiency virus isolates."
RL Arch. Virol. 123:29-45(1992).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; M59418; AAA43074.1; -.
DR PIR; P45557; P45557.
DR InterPro; IPR02050; Env_poliypotein.
DR Pfam; PF00429; Env_poliypotein; 1.
KW Coat protein; Polypeptide; Glycoprotein; Transmembrane.
FT CHAIN 1 610 MAJOR GLYCOPROTEIN GP100.
FT CHAIN 611 855 GLYCOPROTEIN GP36.
FT TRANSMEM 786 809 POTENTIAL.
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 342 342 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 548 548 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 551 551 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 716 716 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 728 728 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 736 736 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 813 813 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 855 AA; 98238 MW; BC2PABBB6245D70D CRC64;

Query Match 68.0%; Score 34; DB 1; Length 855;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```



```

RESULT 11
MAZ4 SCHCO STANDARD; PRT; 940 AA.
ID MAZ4 SCHCO STANDARD; PRT; 940 AA.
AC P37938;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mating-type protein A-alpha 24.
OS Schizophyllum commune (Bracket fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Schizophyllaceae; Schizophyllum.
OX NCBI_TaxId=5334;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CBS 340.81 / UTM 4-40;
RA MEDLINE=92357793; PubMed=135386;
RA Stankis M.M., Specht C.A., Yang H., Glasson L., Ulrich R.C.,
RA Novorny C.P.;
RT "The A alpha mating locus of Schizophyllum commune encodes two
RT dissimilar multiallelic homeobox proteins."
RC Proc. Natl. Acad. Sci. U.S.A. 89:7169-7173(1992).
CC - FUNCTION: SPECIFICS A-ALPHA-4 MATING-TYPE. MAY REGULATE THE
CC EXPRESSION OF GENES SPECIFIC TO THE HOMOKARYOTIC CELL TYPE.
CC - SUBCELLULAR LOCATION: Nuclear.
CC - DEVELOPMENTAL STAGE: Expressed constitutively in homokaryons.
CC - SIMILARITY: Belongs to the TALE/M-ATYP homeobox family.
CC - SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; M97181; AAB01372.1; -.
CC PIR; D37271; D37271.
CC DR InterPro; IPR008422; Coprinus_mating.
CC DR InterPro; IPR001356; Homeobox.
CC DR Pfam; PF05920; Coprinus_mating; 1.
CC DR SMART; SM00389; HOX; 1.
CC DR PROSITE; PS00027; HOMEBOX_1; FALSE NEG.
CC DR PROSITE; PS50071; HOMEBOX_2; FALSE NEG.
CC KM Homeobox; DNA-binding; Transcription regulation; Nuclear protein.
CC FT DNA BIND 110 182 HOMEBOX (TALE-TYPE).
CC FT DOMAIN 370 402 ASH/GLU-RICH (ACIDIC).
CC FT DOMAIN 436 469 ASP/GLU-RICH (ACIDIC).
CC SQ SEQUENCE 940 AA; 101856 MW; 4B99CBAEDB39621E CRC64;

Query Match 68.0%; Score 34; DB 1; Length 940;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAAARVWV 10
DB 172 IAAARWGW 180

RESULT 12
REP2_ECOLI STANDARD; PRT; 285 AA.
ID REP2_ECOLI STANDARD; PRT; 285 AA.
AC P03066; Q47411;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Replication initiation protein.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;

```

```

RN
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII R100;
RX MEDLINE=81074309; PubMed=7003300;
RA Rosen U., Ryder T., Inokuchi H., Ohtsubo H., Ohtsubo E.;
RT "Genes and sites involved in replication and incompatibility of an
RT R100 plasmid derivative based on nucleotide sequence analysis."
RL Mol. Genet. 179:527-537(1980).
RN [2]
RP SEQUENCE OF 1-77 FROM N.A.
RC PLASMID-IncFII R100, and IncFII R1;
RX MEDLINE=81173118; PubMed=6163994;
RA Rosen U., Ryder T., Ohtsubo H., Ohtsubo E.;
RT "Role of RNA transcripts in replication incompatibility and copy
RT number control in antibiotic resistance plasmid derivatives."
RL Nature 290:794-797(1981).
RN [3]
RP SEQUENCE OF 270-285 FROM N.A.
RC PLASMID-IncFII R1;
RX MEDLINE=88289416; PubMed=3041379;
RA Masai H., Arai K.;
RT "Repa protein- and oriR-dependent initiation of R1 plasmid
RT replication: identification of a rho-dependent transcription
RT terminator required for cis-action of repA protein."
RL Nucleic Acids Res. 16:6493-6514(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII R1;
RX MEDLINE=89011975; PubMed=3050127;
RA Dong X., Womble D.D., Rownd R.H.;
RT "In-vivo studies on the cis-acting replication initiator protein of
RT IncFII plasmid NR1."
RL J. Mol. Biol. 202:495-509(1988).
RN [5]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII R1;
RX MEDLINE=85160860; PubMed=2580099;
RA Womble D.D., Sampathkumar P., Easton A.M., Luckow V.A., Rownd R.H.;
RT "Transcription of the replication control region of the IncFII
RT R-plasmid NR1 in vitro and in vivo."
RL J. Mol. Biol. 181:395-410(1985).
RN [6]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII R100;
RX MEDLINE=86319522; PubMed=3019092;
RA Ohtsubo H., Ryder T.B., Maeda Y., Armstrong K., Ohtsubo E.;
RT "DNA replication of the resistance plasmid R100 and its control."
RL Adv. Biophys. 21:115-133(1986).
CC - FUNCTION: This protein is essential for plasmid replication, it is
CC involved in copy control functions.
CC - SIMILARITY: BELONGS TO THE INCFII REPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; J01762; AA92257.1; -.
CC DR EMBL; J01770; ?; NOT ANNOTATED_CDS.
CC DR EMBL; X12587; CAA31100.1; -.
CC DR EMBL; X12776; CAA31263.1; -.
CC DR EMBL; X02302; CAA26168.1; -.
CC DR EMBL; X02302; CAA26169.1; ALT_INIT.
CC DR EMBL; M28840; AAA26067.1; -.
CC DR PIR; A03602; IDECRP.
CC DR PIR; I64780; I64780.
CC DR InterPro; IPR003446; Replicn.
CC DR Pfam; PF02387; IncFII_repa; 1.
CC PLASMID; DNA replication; Plasmid copy control.
CC CONFLICT 55 55 R -> H (IN REF. 6).

```

SEQ SEQUENCE 285 AA; 32755 MW; A21C9D59D24B26B3 CRC64;  
 Query Match 66.0%; Score 33; DB 1; Length 285;  
 Best Local Similarity 75.0%; Pred. No. 25;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AARARVW 10  
 |||||  
 DB 168 ARSRVW 175

RESULT 13  
 SYM LISIN STANDARD; PRT; 331 AA.

AC Q929H5; 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TPRS)  
 GN TRPS OR L1N2301.  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1642;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusnick C., Amend A.,  
 Baquero P., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,  
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,  
 Entian K.-D., Fahl H., Garcia-del Portillo F., Garrido P.,  
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut U., Jackson D.,  
 Jones L.-M., Kaerst U., Kreft U., Kuhn M., Kunst F., Kurapkat G.,  
 Madueno E., Maitounan A., Mata Vicente J., Ng B., Nedjari H.,  
 Nordstrek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Coessart P.;  
 RA "Comparative genomics of Listeria species.";

RT Science 294:849-852(2001).  
 CC -1 CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC -1 SUBUNIT: Homodimer (By similarity).  
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic.

CC -1 SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL, AL596171; CAC97529.1; -  
 DR PIR, A11719; A11719.  
 DR Listlist; L1N2301; -  
 DR HAMAP, MF 00140; -; 1.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR004142; tRNA-synt\_1.  
 DR InterPro; IPR002306; Trp\_tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA tRNA LIGASE I; 1.  
 KW Aminoacyl-tRNA synthetase, Protein biosynthesis, Ligase, ATP-binding;  
 KW Complete proteome.  
 FT SITE 11 19 "HIGH" REGION.  
 FT SITE 193 197 "KMSKS" REGION.  
 FT BINDING 196 196 ATP (By SIMILARITY).  
 SQ SEQUENCE 331 AA; 36975 MW; 6050DC94C9337B73 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 331;  
 Best Local Similarity 60.0%; Pred. No. 29;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVARARVW 10  
 |||||  
 DB 83 EVAAHQAW 92

RESULT 14  
 SYM LISMO STANDARD; PRT; 331 AA.

AC O8Y577; 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TPRS)  
 GN TRPS OR LMO2198.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD-e / Serovar 1/2a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusnick C., Amend A.,  
 Baquero P., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,  
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,  
 Entian K.-D., Fahl H., Garcia-del Portillo F., Garrido P.,  
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut U., Jackson D.,  
 Jones L.-M., Kaerst U., Kreft U., Kuhn M., Kunst F., Kurapkat G.,  
 Madueno E., Maitounan A., Mata Vicente J., Ng B., Nedjari H.,  
 Nordstrek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Coessart P.;  
 RA "Comparative genomics of Listeria species.";

RT Science 294:849-852(2001).  
 CC -1 CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC -1 SUBUNIT: Homodimer (By similarity).  
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic.

CC -1 SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL, AL591982; CAD00276.1; -  
 DR PIR, AF1349; AF1349.  
 DR Listlist; LMO2198; -  
 DR HAMAP, MF 00140; -; 1.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR004142; tRNA-synt\_1.  
 DR InterPro; IPR002306; Trp\_tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA tRNA LIGASE I; 1.  
 KW Aminoacyl-tRNA synthetase, Protein biosynthesis, Ligase, ATP-binding;  
 KW Complete proteome.  
 FT SITE 11 19 "HIGH" REGION.  
 FT SITE 193 197 "KMSKS" REGION.  
 FT BINDING 196 196 ATP (By SIMILARITY).  
 SQ SEQUENCE 331 AA; 36892 MW; 00BC24AF163A3518 CRC64;  
 Query Match 66.0%; Score 33; DB 1; Length 331;  
 Best Local Similarity 60.0%; Pred. No. 29;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVAAARVVM 10  
| | | | |  
| | | | |  
Db 83 EVAHAQAAM 92

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVAAARVVM 10  
| | | | |  
| | | | |  
Db 48 KVAARVVM 57

Search completed: September 1, 2004, 00:10:39  
Job time : 6.56667 secs

RESULT 15  
ID YDK9 SCHPO STANDARD; PRT; 1033 AA.  
AC P8715;  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein C20G8.09c in chromosome I.  
GN SPAC20G8.09c.  
OS Schizosaccharomyces pombe (Fission Yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=972;  
RC MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jacobs K.,  
RA Jones K., Jones S., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Voiclaert G., Aert R., Robben J., Gymnopoulos B.,  
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Leinich H., Reinhardt R., Pohl T.M.,  
RA Egger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rocher M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,  
RT "The genome sequence of Schizosaccharomyces pombe."  
RL Nature 415:871-880(2002).  
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -1- SIMILARITY: Belongs to the UPF0202 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC DR EMBL; Z95334; CAB08603.1; -.  
CC DR PIR; T38131; T38131.  
CC DR GeneDB SPombe; SPAC20G8.09c; -.  
CC DR InterPro; IPR007807; DUF699.  
CC DR Pfam; PF05127; DUF699; 1.  
CC KM Hypothetical protein; ATP-binding; Nuclear protein.  
CC NP\_BIND 282 289 ATP (POTENTIAL).  
CC SEQUENCE 1033 AA; 116463 MW; 8432B313D818135 CRC64;  
Query March 66.0%; Score 33; DB 1; Length 1033;  
Best Local Similarity 60.0%; Pred. No. 86;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2004, 23:56:54 ; Search time 18.5333 Seconds  
(without alignments)  
170.244 Million cell updates/sec

Title: US-09-508-095-8  
Perfect score: 50  
Sequence: 1 EVAAARAYVM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_25:\*  
1: sp archaea:\*  
2: sp bacteria:\*  
3: sp fungi:\*  
4: sp human:\*  
5: sp invertebrate:\*  
6: sp mammal:\*  
7: sp mhc:\*  
8: sp organelle:\*  
9: sp phage:\*  
10: sp plant:\*  
11: sp rodent:\*  
12: sp virus:\*  
13: sp vertebrate:\*  
14: sp unclassified:\*  
15: sp virus:\*  
16: sp bacteriophage:\*  
17: sp archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	82.0	980	17 Q8U0T0	Q8U0T0 pyrococcus
2	39.5	79.0	711	4 Q8TCD2	Q8TCD2 homo sapien
3	39.5	79.0	711	4 Q8I2H6	Q8I2H6 homo sapien
4	39.5	79.0	711	4 Q8I2H2	Q8I2H2 homo sapien
5	39	78.0	108	10 Q8H321	Q8H321 oryza sativ
6	39	78.0	175	9 Q38153	Q38153 bacteriophag
7	38	76.0	166	16 Q7VMA7	Q7VMA7 bordetella
8	38	76.0	191	17 Q9VOH2	Q9VOH2 pyrococcus
9	37	74.0	252	16 Q83H01	Q83H01 tripheryma
10	37	74.0	291	16 Q83GH9	Q83GH9 tripheryma
11	37	74.0	372	2 Q83WV1	Q83WV1 thicapsa r
12	36	72.0	181	5 Q86E95	Q86E95 schistosoma
13	36	72.0	365	2 Q7X365	Q7X365 uncultured
14	36	72.0	408	16 Q7VD97	Q7VD97 prochloroco
15	36	72.0	473	16 Q7V935	Q7V935 prochloroco
16	36	72.0	1919	6 Q29518	Q29518 oryctolagus

17	36	72.0	1920	6 Q29519	Q29519 oryctolagus
18	35	70.0	108	17 Q9YA3	Q9YA3 aeropyrum p
19	35	70.0	166	16 Q7WGY9	Q7WGY9 bordetella
20	35	70.0	166	16 Q7W9R9	Q7W9R9 bordetella
21	35	70.0	210	16 Q9A6X8	Q9A6X8 caulobacter
22	35	70.0	254	16 Q82HU3	Q82HU3 streptomyce
23	35	70.0	269	16 Q8R8C1	Q8R8C1 thermomaneer
24	35	70.0	273	2 Q8RPF8	Q8RPF8 rhizobium l
25	35	70.0	309	16 Q8PQ53	Q8PQ53 xanthomonas
26	35	70.0	380	16 Q8UFU2	Q8UFU2 agrobacteri
27	35	70.0	441	16 Q82BH1	Q82BH1 streptomyce
28	35	70.0	689	16 Q92XJ5	Q92XJ5 rhizobium m
29	35	70.0	721	11 Q8BP06	Q8BP06 mus musculu
30	35	70.0	741	11 Q8BP04	Q8BP04 mus musculu
31	35	70.0	1114	16 Q9XYF4	Q9XYF4 streptomyce
32	35	70.0	1664	13 Q8JTF9	Q8JTF9 acanthogobi
33	34	68.0	117	9 Q38071	Q38071 bacteriophag
34	34	68.0	146	16 Q9ZBN2	Q9ZBN2 streptomyce
35	34	68.0	158	10 Q9SWP3	Q9SWP3 arabidopsis
36	34	68.0	213	16 Q9ABY8	Q9ABY8 caulobacter
37	34	68.0	284	16 Q890H7	Q890H7 lactobacilli
38	34	68.0	328	2 Q83ZS0	Q83ZS0 corynebacte
39	34	68.0	387	16 Q66630	Q66630 aquifex aeo
40	34	68.0	390	16 P74336	P74336 synechocyst
41	34	68.0	416	16 Q8P966	Q8P966 xanthomonas
42	34	68.0	430	16 Q82QU8	Q82QU8 streptomyce
43	34	68.0	489	16 Q7UKJ5	Q7UKJ5 rhodospirilli
44	34	68.0	706	16 Q9KR79	Q9KR79 vibrio chol
45	34	68.0	707	11 Q8CBA0	Q8CBA0 mus musculu

#### ALIGNMENTS

RESULT 1	ID	Q8U0T0	PRELIMINARY;	PRT;	980 AA.
AC	Q8U0T0	01-TUN-2002 (TREMblrel. 21, Created)			
DT	01-TUN-2002 (TREMblrel. 21, Last sequence update)				
DT	01-OCT-2003 (TREMblrel. 25, Last annotation update)				
DE	Large helicase-related protein.				
GN	PF1504.				
OS	Pyrococcus furiosus.				
OC	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;				
CC	Pyrococcus.				
OX	NCBI_TaxID=2261;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;				
RA	Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;				
RT	"The complete sequence of the Pyrococcus furiosus genome."				
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AE010251; AAL81628.1				
DR	GO; GO:0005524; F:ATP binding; IEA.				
DR	GO; GO:0008026; F:ATP dependent helicase activity; IEA.				
DR	GO; GO:003676; F:nucleic acid binding; IEA.				
DR	InterPro; IPR001410; DEAD.				
DR	InterPro; IPR001650; Helicase_C.				
DR	Pfam; PF00271; helicase_C; 1.				
DR	SMART; SM00487; DEHDC; 1.				
DR	SMART; SM00490; HELIC; 1.				
DR	Helicase; Complete proteome.				
SQ	SEQUENCE 980 AA; 11258 MW; F244EF8F681C653 CRC64;				
Query Match	82.0%;	Score 41;	DB 17;	Length 980;	
Best Local Similarity	80.0%;	Pred. No. 24;			
Matches	8;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
QY	1 EVAAARAYVM 10				
Db	298 EVATRLAYVM 307				

## RESULT 2

08TCD2 PRELIMINARY; PRT; 711 AA.  
ID 08TCD2  
AC 08TCD2;  
DT 01-JUN-2002 (TREMblrel. 21, Created)  
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
DE Lactoferrin.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strassberg R;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
DR EMBL: BC022347; AAN2347.1; -;  
DR GO: GO:000576; C:extracellular; IEA.  
DR GO: GO:0008199; F:ferric iron binding; IEA.  
DR GO: GO:0006879; P:iron ion homeostasis; IEA.  
DR GO: GO:0006826; P:iron ion transport; IEA.  
DR GO: GO:0006810; P:transport; IEA.  
DR InterPro: IPR001156; Transferrin.  
DR Pfam: PF00405; Transferrin; 2.  
DR PRINTS: PRO0422; TRANSFERRIN.  
DR SMART: SM00094; TR\_FER; 2.  
DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE: PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE: PS00207; TRANSFERRIN\_3; 2.  
KM Glycophenol; Iron transport; Metal-binding; Transport.  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;  
Query Match 79.0%; Score 39.5; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 34;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
OY 1 EVAA-RARYVM 10  
DB 357 EVAARRARVVM 367  
RESULT 3  
08IZH6 PRELIMINARY; PRT; 711 AA.  
AC 08IZH6;  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
DE Lactoferrin.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kaplan J.B.; Fine D.H.;  
RT "Characterization of an amino acid polymorphism in the antibacterial  
RT domain of human lactoferrin."  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY137470; AAN1304.1;  
DR GO: GO:000576; C:extracellular; IEA.  
DR GO: GO:0008199; F:ferric iron binding; IEA.  
DR GO: GO:0006879; P:iron ion homeostasis; IEA.  
DR GO: GO:0006826; P:iron ion transport; IEA.  
DR InterPro: IPR001156; Transferrin.  
DR Pfam: PF00405; Transferrin; 2.

DR PRINTS: PR00422; TRANSFERRIN.  
DR SMART: SM00094; TR\_FER; 2.  
DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE: PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE: PS00207; TRANSFERRIN\_3; 2.  
SQ SEQUENCE 711 AA; 78396 MW; 547AB9423C27CB67 CRC64;  
Query Match 79.0%; Score 39.5; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 34;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
OY 1 EVAA-RARYVM 10  
DB 357 EVAARRARVVM 367

## RESULT 4

08IU92 PRELIMINARY; PRT; 711 AA.  
ID 08IU92  
AC 08IU92;  
DT 01-MAR-2003 (TREMblrel. 23, Created)  
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
DE Lactoferrin precursor (lactoferrin).  
GN LTF.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Shi Y.-Q.; Zhang Y.; Zheng Y.-M.;  
RT "Homo sapiens Lactoferrin Gene: cDNA Cloning and Sequence  
RT Analysis."  
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Semenal vesicle;  
RA Bakker Singh S.; Saravanan K.; Paramasivam M.; Srinivasan A.;  
RT "Homo sapiens Lactoferrin (LTF) mRNA."  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY165046; AAN6398.1; -;  
DR EMBL: AY178998; AAN75578.2; -;  
DR GO: GO:000576; C:extracellular; IEA.  
DR GO: GO:0008199; F:ferric iron binding; IEA.  
DR GO: GO:0006879; P:iron ion homeostasis; IEA.  
DR GO: GO:0006826; P:iron ion transport; IEA.  
DR InterPro: IPR001156; Transferrin.  
DR Pfam: PF00405; Transferrin; 2.  
DR PRINTS: PRO0422; TRANSFERRIN.  
DR SMART: SM00094; TR\_FER; 2.  
DR PROSITE: PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE: PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE: PS00207; TRANSFERRIN\_3; 2.  
KM Signal.  
FT SIGNAL.  
FT CHAIN 1 19 POTENTIAL.  
FT CHAIN 20 711 LACTOFERRIN.  
SQ SEQUENCE 711 AA; 78382 MW; 547BFC4C9267E67 CRC64;  
Query Match 79.0%; Score 39.5; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 34;  
Matches 10; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
OY 1 EVAA-RARYVM 10  
DB 357 EVAARRARVVM 367

## RESULT 5

08H321 PRELIMINARY; PRT; 108 AA.  
ID 08H321

AC OGB321;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE OSJNBA0077M12.27 protein.  
 GN OSJNBA0077M12.27.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC  
 clone:OSJNBA0077M12.27";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005301; BAC24969.1; -  
 SO SEQUENCE 108 AA; 10826 MW; 75A593DEF4ECC953 CRC64;

Query Match 78.0%; Score 39; DB 10; Length 108;  
 Best Local Similarity 80.0%; Pred. No. 6.5;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVAARARVW 10  
 Db 62 EAAARARVW 71

RESULT 6  
 ID Q38153 PRELIMINARY; PRT; 175 AA.  
 AC Q38153;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE ORF4.  
 OS Bacteriophage SPPL.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
 OC Lambda-like viruses.  
 OX NCBI\_TaxID=10724;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94172631; PubMed=8126723;  
 RA Pedre X., Weise F., Chai S., Lueder G., Alonso J.C.;  
 RT "Analysis of cis and trans acting elements required for the initiation  
 of DNA replication in the Bacillus subtilis bacteriophage SPPL.";  
 RL J. Mol. Biol. 236:1324-1340(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Alonso J.C., Luder G., Stiege A.C., Chai S., Weise F., Trautner T.A.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Alonso J.C.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X67865; CAA48064.1; -  
 DR EMBL; X97918; CAA6499.1; -  
 DR PIR; S43815; S43815.  
 SO SEQUENCE 175 AA; 20106 MW; 2B4B409C636055C3 CRC64;

Query Match 78.0%; Score 39; DB 9; Length 175;  
 Best Local Similarity 60.0%; Pred. No. 11;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVAARARVW 10  
 Db 106 EISAKRVRW 115

RESULT 7  
 Q7VWA7

ID Q7VWA7 PRELIMINARY; PRT; 166 AA.  
 AC Q7VWA7;  
 DT 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Putative inner membrane transport protein.  
 GN BP2353.  
 OS Bordetella pertussis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=520;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tohama 1 / ATCC BAA-589 / NCTC 13251;  
 RX MEDLINE=22827954; PubMed=12910271;  
 RA Parkhill J., Sebatina M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Farraga A.M., Temple L., James K., Barton N., Cherevach I.,  
 RA Achtman M., Aikin R., Baker S., Basham D., Bason N., Doggett J.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Duggett J.,  
 RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagsels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640418; CAA42626.1; -  
 KW Complete proteome.  
 SO SEQUENCE 166 AA; 18251 MW; E5E685E71F0C416 CRC64;

Query Match 76.0%; Score 38; DB 16; Length 166;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVAARARVW 10  
 Db 48 EVATRAQLW 57

RESULT 8  
 ID Q9V0H2 PRELIMINARY; PRT; 912 AA.  
 AC Q9V0H2;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Large helicase-related protein (LHR-1).  
 GN PYRAB08170 OR PAB1517.  
 OS Pyrococcus abyssi.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GES / Orsay;  
 RA Hellig R.;  
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome  
 structure and evolution.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO DEAD/DEAH BOX HELICASE FAMILY.  
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.  
 DR EMBL; AJ248285; CAA9731.1; -  
 DR PIR; B75127; B75127.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase C; 1.  
 DR SMART; SM00487; DEADc; 1.

DR SMART; SM00490; HELIC; 1.  
 KM ATP-binding; Helicase; Complete proteome.  
 SQ SEQUENCE 912 AA; 104689 MW; 36f845062D01B811 CRC64;

Query Match 76.0%; Score 38; DB 17; Length 912;  
 Best Local Similarity 70.0%; Pred. No. 87;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVARARVW 10  
 DB 230 EVARARLW 239

## RESULT 9

Q83H01 PRELIMINARY; PRT; 252 AA.

AC Q83H01 01-JUN-2003 (TREMblrel. 24, Created)  
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
 DE 6-phosphogluconolactonase (EC 3.1.1.31).  
 GN PGL OR TW464.  
 OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Micrococcineae; Cellulomonadaceae; Tropheryma.  
 CX NCBI\_TaxID=218496;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22495039; PubMed=12606174;  
 RA Bentley S.D., Maiwald M., Murphy L.D., Pallen M.J., Yeats C.A.,  
 RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,  
 RA von Heijay A., Goble A., Rutter S., Squares R., Squares S.,  
 RA Barrell B.G., Parkhill J., Rellman D.A.,  
 RT "Sequencing and analysis of the genome of the Whipple's disease  
 bacterium Tropheryma whipplei.";  
 RT Lancet 361:637-644(2003).  
 DR EMBL; BX251411; CAD67132.1; -  
 DR GO; GO:0017057; F:6-phosphogluconolactonase activity; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.  
 DR InterPro; IPR006148; Gluc\_gal\_isom.  
 DR Pfam; PF01182; Glucosamine Iso; 1.  
 DR TIGRFAMs; TIGR01198; pgl; 1.  
 KM Hydroxylase; Complete proteome.  
 SQ SEQUENCE 252 AA; 28115 MW; 02046BE1FDC2101 CRC64;

Query Match 74.0%; Score 37; DB 16; Length 252;  
 Best Local Similarity 75.0%; Pred. No. 38;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AARARVW 10  
 DB 60 AARARVW 67

## RESULT 10

Q83GH9 PRELIMINARY; PRT; 291 AA.

AC Q83GH9 01-JUN-2003 (TREMblrel. 24, Created)  
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
 DE 6-phosphogluconolactonase (EC 3.1.1.31).  
 GN DEVA OR TW1308.  
 OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Micrococcineae; Cellulomonadaceae; Tropheryma.  
 CX NCBI\_TaxID=203267;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Raulit D., Audic S., Robert C., Ogata H., Suhre K., Drancourt M.,

RA Claverie J.-M.;  
 RT "Tropheryma whipplei illustrates the diversity of gene loss patterns  
 in small genome bacterial pathogens.";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE016851; AA04405.1; -  
 DR GO; GO:0017057; F:6-phosphogluconolactonase activity; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.  
 DR InterPro; IPR006148; Gluc\_gal\_isom.  
 DR InterPro; IPR005900; Phosphogluconac.  
 DR Pfam; PF01182; Glucosamine Iso; 1.  
 DR TIGRFAMs; TIGR01198; pgl; 1.  
 KM Hydroxylase; Complete proteome.  
 SQ SEQUENCE 291 AA; 32564 MW; FC80C8A855EB1B CRC64;

Query Match 74.0%; Score 37; DB 16; Length 291;  
 Best Local Similarity 75.0%; Pred. No. 44;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AARARVW 10  
 DB 99 AARARVW 106

## RESULT 11

Q83WV1 PRELIMINARY; PRT; 372 AA.

AC Q83WV1 01-JUN-2003 (TREMblrel. 24, Created)  
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
 DE ENVZ (Fragment).  
 GN ENVZ.  
 OS Thiacapsa roseopersicina.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;  
 OC Chromatiaceae; Thiacapsa.  
 CX NCBI\_TaxID=1058;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BBS;  
 RA Marotti G., Fodor B.D., Rakhely G., Kovacs A.T., Arvany S.,  
 RA Kovacs K.L.;  
 RT "Accessory proteins functioning selectively and pleiotropically in the  
 biosynthesis of [NFe] hydrogenases in Thiacapsa roseopersicina.";  
 RL Eur. J. Biochem. 270:2218-2227(2003).  
 DR EMBL; AY152822; AAN87038.1; -

DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0016301; F:kinase activity; IEA.  
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR003660; HAMP.  
 DR InterPro; IPR005467; His\_kinase.  
 DR InterPro; IPR003661; His\_kin\_N.  
 DR Pfam; PR00672; HAMP; 1.  
 DR Pfam; PR00512; HisKA; 1.  
 DR SMART; SW00304; HAMP; 1.  
 DR SMART; SW00388; HisKA; 1.  
 DR PROSITE; PSS0885; HAMP; 1.  
 DR PROSITE; PSS0109; HIS\_KIN; 1.  
 FT NON TER 372  
 SQ SEQUENCE 372 AA; 41488 MW; 90BC56338709EDD CRC64;

Query Match 74.0%; Score 37; DB 2; Length 372;  
 Best Local Similarity 66.7%; Pred. No. 56;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VAARARVW 10  
 DB 308 VAARARVW 316

## RESULT 12



```

Q86E95
ID Q86E95 PRELIMINARY; PRT; 161 AA.
AC Q86E95;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Clonase Z22675 mRNA sequence.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu F., Hu W., Yan Q., Xu X., Zhu Z., Zhang X., Wang J., Sun J.,
RA Xu X., Wang Z., Zeng L., Rong Y., Wu X., Qu J., Xu Z., Huang J.,
RA Ma Y., Wang S., Wang Z., Xue C., Feng Z., Chen Z., Han Z.;
RT "The full-length cDNA of S. japonicum genes."
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY223325; AAP06362.1; -.
SQ SEQUENCE 161 AA; 21357 MW; F2371079F65AE25A CRC64;

Query Match 72.0%; Score 36; DB 5; Length 161;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVARARVW 10
Db 103 EVSGRYRIW 112

RESULT 13
Q7X365 PRELIMINARY; PRT; 365 AA.
ID Q7X365;
AC Q7X365;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative transport protein.
OS uncultured Acidobacteria bacterium.
OC Bacteria; Acidobacteria; environmental samples.
OX NCBI_TaxID=171953;
RN [1]
RP SEQUENCE FROM N.A.
RA Quaiser A., Ochsenreiter T., Lanz C., Schuster S.C., Treusch A.H.,
RA Bok J., Schleper C.;
RT "Acidobacteria form a coherent but highly diverse group within the
RT bacterial domain and exchange genes with highly diverse group within the
RT population genomes."
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY281352; AAP58499.1; -.
SQ SEQUENCE 365 AA; 38767 MW; 07C6AB9231633C08 CRC64;

Query Match 72.0%; Score 36; DB 2; Length 365;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 ARARVW 10
Db 281 ARARVW 287

RESULT 14
Q7VD97 PRELIMINARY; PRT; 408 AA.
ID Q7VD97;
AC Q7VD97;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Uncharacterized conserved protein.
GN PRO0486.
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.

```

```

OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SARG / CCMF 1375 / SS120;
RC MEDLINE=22810154; PubMed=12917486;
RA Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
RA Barre V., Duprat S., Gaipierin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Ozta S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissendbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.;
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal oxyphototrophic genome."
DR Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
KW Complete proteome.
SQ SEQUENCE 408 AA; 45243 MW; 772AEFD0BEA1A9B CRC64;

Query Match 72.0%; Score 36; DB 16; Length 408;
Best Local Similarity 50.0%; Pred. No. 96;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVARARVW 10
Db 313 ELAAKVRRIW 322

RESULT 15
Q7V935 PRELIMINARY; PRT; 473 AA.
ID Q7V935;
AC Q7V935;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Mgte family, putative magnesium transport protein (BC 3.6.1.-).
GN PMT0126.
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser J., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Stieglitz C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation."
RL Nature 424:1042-1047(2003).
DR EMBL: BX572095; CAE20301.1; -.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 473 AA; 51395 MW; 7DCBE9DB78DD673 CRC64;

Query Match 72.0%; Score 36; DB 16; Length 473;
Best Local Similarity 77.8%; Pred. No. 11e02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VARARVW 10
Db 297 VVARVW 305

Search completed: September 1, 2004, 00:15:26
Job time : 24.5333 secs

```

